TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

EINAR HILLE C. C. MACDUFFEE OSCAR ZARISKI

WITH THE COÖPERATION OF

A. A. ALBERT
T. H. HILDEBRANDT
OYSTEIN ORE
J. L. SYNGE
JESSE DOUGLAS
R. E. LANGER
H. L. RIETZ
GABOR SZEGÖ
R. L. WILDER
NELSON DUNFORD
SAUNDERS MACLANE
H. P. ROBERTSON
G. T. WHYBURN

VOLUME 50

JULY TO DECEMBER, 1941

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1941

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

Chester C. Corbin Library Fund
Bound Dec. 1941

Composed, Printed and Bound by
The Collegiate Press
George Banta Publishing Company
Menasha, Wisconsin

67382

QA
1
F001
v. 50

TABLE OF CONTENTS

VOLUME 50, JULY TO DECEMBER, 1941

BELL, P. O. On differential geometry intrinsically connected with a surface element of projective arc length 529
CLIFFORD, A. H., and MACLANE, S. Factor-sets of a group in its abstract unit group 385
COBURN, N. Conformal unitary spaces 26
COURANT, R. On a generalized form of Plateau's problem 40
CURRY, H. B. The paradox of Kleene and Rosser 454
DOUGLAS, J. Solution of the inverse problem of the calculus of variations 71
DUFFIN, R. J., and SCHAEFFER, A. C. A refinement of an inequality of the brothers Markoff 517
JACOBSON, N. Restricted Lie algebras of characteristic p 15
JENNINGS, S. A. The structure of the group ring of a p-group over a modular field 175
KIOKEMEISTER, F. The parastrophic criterion for the factorization of primes 140
KOBER, H. On a theorem of Schur and on fractional integrals of purely imaginary order 160
LOOMIS, L. H. The decomposition of meromorphic functions into rational functions of univalent functions 1
MACLANE, S., and CLIFFORD, A. H. Factor-sets of a group in its abstract unit group 385
MACLANE, S., and SCHILLING, O. F. G. Normal algebraic number fields. 295
MARCINKIEWICZ, J., and ZYGMUND, A. On the behavior of trigonometric series and power series 407
VON NEUMANN, J., and SCHOENBERG, I. J. Fourier integrals and metric geometry 226
PÓLYA, G. Sur l'existence de fonctions entières satisfaisant à certaines conditions linéaires 129
SCHAEFFER, A. C., and DUFFIN, R. J. A refinement of an inequality of the brothers Markoff 517
SCHAEFFER, A. C., and SZEGÖ, G. Inequalities for harmonic polynomials in two and three dimensions 187
SCHILLING, O. F. G., and MACLANE, S. Normal algebraic number fields. 295
SCHOENBERG, I. J., and VON NEUMANN, J. Fourier integrals and metric geometry 226
SZEGÖ, G., and SCHAEFFER, A. C. Inequalities for harmonic polynomials in two and three dimensions 187

TRJITZINSKY, W. J. Properties of growth for solutions of differential equations of dynamical type 252
ZARISKI, O. Pencils on an algebraic variety and a new proof of a theorem of Bertini 48
ZYGMUND, A., and MARCINKIEWICZ, J. On the behavior of trigonometric series and power series 407

THE DECOMPOSITION OF MEROMORPHIC FUNCTIONS INTO RATIONAL FUNCTIONS OF UNIVALENT FUNCTIONS

BY

LYNN H. LOOMIS[1]

1. **Introduction.** If $f(z)$ is meromorphic at a point z_0 then it is an elementary fact that the multivalent character of $f(z)$ in the neighborhood of z_0 is like that of a power. By this we mean that for some integer n, the function $[f(z)-f(z_0)]^{1/n}$ (or $f(z)^{-1/n}$ if z_0 is a pole; for simplicity we assume $f(z)$ to be regular at z_0) is univalent in a sufficiently small circle $|z-z_0|<r$. This property is clearly a local property. Let us consider other ways of expressing it. We may say that the map of a sufficiently small circle $|z-z_0|<r$ under $w=f(z)$ is a Riemann configuration[2] which can be imbedded in the Riemann surface of $[w-f(z_0)]^{1/n}$. Or we may say that $f(z)$ has a decomposition $f(z)\equiv f_2(f_1(z))$, where $\zeta=f_1(z)$ is univalent in a sufficiently small circle $|z-z_0|<r$, and $f_2(\zeta)$ is the rational function $\zeta^n+f(z_0)$.

It is the object of this paper to study the analogue in the large of the local property just described. That is, given $f(z)$ meromorphic in a simply-connected region R, we shall investigate the decomposition $f(z)\equiv f_2(f_1(z))$, where $f_2(\zeta)$ is a rational function, and $\zeta=f_1(z)$ is univalent in R. In particular we shall answer the following questions. If $f(z)$ is meromorphic in $|z|<1$, what are necessary and sufficient conditions that $f(z)$ be decomposable as above? What modifications of these conditions are equivalent to restricting $\zeta=f_1(z)$ to mapping functions for special classes of simply-connected regions, say Jordan regions? What modifications are equivalent to restricting $f_2(\zeta)$ to special classes of rational functions, say polynomials?

2. **Survey.** In this section we shall attempt to motivate the line of development pursued and indicate the nature of the theorems to be proved, as well as introduce whatever terminology will be needed and state what facts will be assumed.

The general question which we have raised is this: given $f(z)$ meromorphic in a simply-connected region R of the z-plane, what conditions are necessary and sufficient that $f(z)$ can be decomposed, $f(z)\equiv f_2(f_1(z))$, where $f_2(\zeta)$ is a rational function, and $\zeta=f_1(z)$ is univalent in R? It is evidently sufficient to discuss the case where R is the unit circle $|z|<1$. For if R has only one boundary

Presented to the Society, October 26, 1940; received by the editors June 25, 1940.

[1] Society of Fellows, Harvard University.

[2] The term *Riemann surface* is often taken to refer to the complete Riemann surface of a given analytic function. The term *Riemann configuration* is used to include both Riemann surfaces in the above sense and pieces of Riemann surfaces.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

point, or is the extended z-plane, then $f_1(z)$ must be linear and $f(z)$ must itself be a rational function. Excluding this trivial case, R has at least two boundary points and can be mapped onto the unit circle $|z|<1$. And the combination of the mapping function with $f_1(z)$ gives another univalent function.

Now let us consider the properties of a function which has such a decomposition. In order to do this we shall need a notion of neighborhood for boundary points. By a *half-neighborhood* of a point z_0 on the boundary $|z|=1$ we shall mean the intersection of the unit circle $|z|<1$ with a circle $|z-z_0|<r$. Note that a half-neighborhood of z_0 does not contain z_0. By a *rational surface* we shall mean the Riemann surface of the inverse of a rational function. Now if $f(z)$ is decomposable, it is trivially true that for every z_0 on $|z|=1$ there is a half-neighborhood whose map under $w=f(z)$ is a Riemann configuration which can be imbedded in a rational surface. In our fundamental (and most general) decomposition theorem, proved in §3, this necessary condition will be proved to be sufficient.

Along the circumference, however, the multivalent character of $f(z)$ is not in general of the simple type occurring at interior points, where $f(z)$ is like a power ζ^n+c. Consider for example a plane simply-connected region in the ζ-plane, of which not all boundary points are attainable. Then if $\zeta=f_1(z)$ maps the unit circle $|z|<1$ onto this region, there is at least one point z_0 on $|z|=1$ where $f_1(z)$ has more than one limiting value. Let $w=f_2(\zeta)$ be a rational function having zeros of $f_2'(\zeta)$ at two such limiting values. Then $f_2(f_1(z))$ is a decomposable function, but its multivalent character at z_0 is not like that of a power. We shall say that $f(z)$ is *rationally-multivalent* at z_0 if there is a neighborhood (half-neighborhood if z_0 lies on the boundary) of z_0 which $w=f(z)$ maps onto a Riemann configuration which can be imbedded in a rational surface (i.e., if there is a rational function $w=f_2(\zeta)$ such that $f_2^{-1}(f(z))$ is meromorphic and univalent in a sufficiently small neighborhood of z_0). If the Riemann surface of an nth root is sufficient for this imbedding (as it is, for instance, at all points of $|z|<1$) we shall say that $f(z)$ is *power-multivalent* at z_0.

Now let us consider the decomposition of $f(z)$ with respect to each of the two functions into which $f(z)$ is decomposed. The function $\zeta=f_1(z)$ is univalent on $|z|<1$ and is therefore a mapping function for a simply-connected region R of the ζ-plane. Now in the theory of conformal mapping, simply-connected regions are often classified according to how much is assumed about the nature of the boundary, the classification running as follows: general simply-connected regions (prime end theory), simply-connected regions all of whose boundary points are attainable, and Jordan regions[3]. The corresponding mapping functions are respectively univalent in $|z|<1$, continuous on $|z|\leqq 1$ and univalent in $|z|<1$, and continuous and univalent in $|z|\leqq 1$. We shall

[3] A *Jordan region* is a region bounded by a Jordan curve. A *Jordan curve* is a simple closed curve; i.e., a one-to-one continuous map of a circle. A *Jordan arc* is a continuous one-to-one map of the closed unit interval $[0, 1]$.

deduce a different form for the decomposition theorem corresponding to each of these classes of regions (or classes of mapping functions); and also a form corresponding to the class of Jordan regions bounded by analytic curves (the class of mapping functions $\zeta = f_1(z)$ regular and univalent on $|z| \leqq 1$). Two of these theorems are direct corollaries of the theorem for the most general case, and will be treated with that theorem in §3. But the theorem corresponding to Jordan regions is deeper and requires considerable development.

This development is suggested by the rather remarkable theorem (which together with related theorems, we shall discuss in §4) that if $f(z)$ is regular in $|z| < 1$, and if the limiting values of $f(z)$ as $|z| \to 1$ all lie on a Jordan curve Γ of the w-plane, then $f(z)$ can be decomposed as $f(z) \equiv f_2(f_1(z))$, where $\zeta = f_1(z)$ is a rational function of the form

$$f_1(z) \equiv \prod_{j=1}^{n} \frac{z - a_j}{1 - \bar{a}_j z}, \qquad |a_j| < 1,$$

and $f_2(\zeta)$ is a mapping function for the inner Jordan region bounded by Γ. The rational functions of the form

$$w = f(z) \equiv K \prod_{j=1}^{n} \frac{z - a_j}{1 - \bar{a}_j z}, \qquad |a_j| < 1,$$

will occur often in this paper, and we shall call them *circular rational functions*. They are the functions(4) which map the unit circle $|z| \leqq 1$ onto the circle $|w| \leqq |K|$ with each value taken precisely n times.

One result of the above theorem is that $f(z)$ is continuous on the closed circle $|z| \leqq 1$, and a second result is that $f(z)$ is power-multivalent (in fact, locally univalent) along $|z| = 1$. These conclusions suggest the following possible generalizations. First, if $f(z)$ is meromorphic in $|z| < 1$, and if the limiting values of $f(z)$ as z approaches the points of an arc γ of $|z| = 1$ lie on a Jordan arc Γ of the w-plane, then $f(z)$ is continuous on γ. Second, under the same circumstances, if $f(z)$ is continuous on γ then $f(z)$ is power-multivalent along γ. In §5 we shall prove both of these statements, and from them, together with an earlier theorem, we shall have at once our best decomposition theorem (corresponding to the class of Jordan regions).

We have not yet considered the rational function $f_2(\zeta)$ in any detail. It is fairly evident that in the cases considered so far (i.e., those involving $f(z)$ meromorphic on $|z| < 1$), nothing can be inferred about the nature of $f_2(\zeta)$ beyond the fact that it is rational. But if $f(z)$ is regular and bounded on $|z| < 1$, then certain of the decomposition theorems can be proved for restricted classes of rational functions. To be explicit, all of the theorems whose premises include (perhaps implicitly) the continuity of $f(z)$ on $|z| \leqq 1$ can be

(4) See, for example, Julia, *Principes Géométriques d'Analyse*, vol. 1, p. 54.

proved with $f_2(\zeta)$ restricted either to polynomials or to circular rational functions. The general theorem, however, cannot be proved in this case, as we shall show by an example. These theorems and this example will be taken up in §6.

Before proceeding to the first theorem let us recall a few function-theoretic facts.

First, the surfaces which we have called rational surfaces are precisely the surfaces composed of n extended planes joined around branch points whose total order is $2n-2$. For every rational surface has this construction(5). And conversely, any such surface is simply-connected (having an Euler-Poincaré characteristic of -2 as any subdivision will show)(6), and hence can be mapped by the fundamental mapping theorem onto a normal region(7), which must in this case be the extended plane. But any function meromorphic on the extended plane is rational.

Second, the Riemann configurations onto which the circular rational functions map the unit circle $|z| \leqq 1$ (we shall call them surfaces of type II) are precisely those configurations composed of n circular discs $|w| \leqq M$ joined around branch points whose total order is $n-1$. For every surface of type II has this construction(8). And conversely, any such surface is simply-connected (with characteristic -1)(9) and hence mappable onto a normal region which in this case must be a circle. But any function which maps the unit circle onto $|w| < M$ with each value taken precisely n times is a circular rational function(10).

We shall also assume one or two results centering around the theorem of Fatou(11) that if $f(z)$ is regular and bounded on $|z| < 1$, then the radial limit $\lim_{r \to 1} f(re^{i\theta})$ exists for almost all θ. A direct corollary affirms the same result for $f(z)$ univalent on $|z| < 1$. Also in this connection we note the theorem of F. and M. Riesz(12) that if $f(z)$ is regular and bounded and not identically constant, then the points of $|z| = 1$ where the radial limit is zero form a set of Lebesgue measure zero. Finally, by a theorem of Lindelöf(13), if $f(z)$ is

---

(5) See, for example, Julia, loc. cit., p. 60.

(6) See Kerékjártó, *Vorlesungen über Topologie*, I, pp. 143–147.

(7) The difficult fundamental theorem of uniformization is not actually needed here. It is sufficient to divide the surface in question into a finite number of plane pieces, and apply the theorem on the blending of the maps of contiguous or overlapping regions. See, for example, Bieberbach, *Funktionentheorie*, vol. 2, p. 171.

(8) See Julia, loc. cit., pp. 54–56, 60. The rational surface of a circular rational function is symmetric with respect to inversion in the circle $|w| = |K|$, and any rational surface of n sheets has branch points of total order $2n-2$.

(9) See Footnote 6.

(10) See Footnote 4.

(11) See Bieberbach, loc. cit., pp. 147–155.

(12) See Bieberbach, loc. cit., pp. 155–156.

(13) See Bieberbach, loc. cit., p. 21.

regular and bounded on $|z|<1$ and if $\lim_{z\to z_0} f(z)$ exists along two Jordan arcs lying in $|z|<1$ except for the common endpoint z_0, then the two limits are the same and $f(z)$ uniformly approaches the limit between the two arcs.

A Jordan arc which lies in a region R except for a single endpoint which lies on the boundary of R is known as a *Jordan cut* of R. A Jordan arc which lies in R except for both endpoints which lie on the boundary of R is known as a Jordan *crosscut* of R. On a general Riemann configuration W a Jordan cut will be defined as a semi-open Jordan arc, whose open end converges to the boundary of W in such a way that its projection on the extended w-plane converges to a point.

We need to make one more general remark. As long as we consider meromorphic functions, such notions as *limit*, *continuity*, *Jordan arc*, etc., must be extended to allow infinite values. This extension is the usual one, geometrized by projecting to the Riemann sphere, and considering limits, continuity, Jordan arcs, and so on, in the ordinary sense for the corresponding sets of points there.

3. **The fundamental decomposition theorem.** We prove first a lemma.

LEMMA 1. *If a Riemann configuration W can be divided by a Jordan crosscut into two Riemann configurations each of which can be imbedded in the Riemann surface of the inverse of a rational function, then W itself can be so imbedded.*

Proof. Let γ, with endpoints A and B, be the crosscut dividing W into W_1 and W_2, and let W_1 and W_2 be imbedded in rational surfaces R_1 and R_2 respectively. Cut each of R_1 and R_2 along γ, and join the left of γ on R_1 to the right of γ on R_2, and vice versa. (Here left and right are used simply to designate the two sides of γ.)

It is obvious that the above construction can change branch point order only on γ. Let p and q be the orders of the branch points at A of R_1 and R_2 respectively, with the convention that, for instance, $p=0$ if R_1 has no branch point at A. Then R_1 and R_2 have respectively $p+1$ and $q+1$ sheets joined around A. The new surface R has $p+q+2$ sheets joined around A, i.e., has a branch of order $p+q+1$ at A. Now consider an interior point C of γ. The construction given carries the two copies γ_1 and γ_2 of γ on R_1 and R_2 into two copies forming a closed curve on R. The two points C_1 and C_2 on γ_1 and γ_2 correspond to distinct points of R. Hence the total order of the two branch points remains the same. For, although the branch points at both C_1 and C_2 may be changed in the construction, the total number of sheets joined around these two points does not change. Thus the method of joining R_1 and R_2 defined above has increased by 2 the total order of the branch points over that for the two surfaces separately (an increase of 1 at each of A and B).

Let the number of sheets of R_1 and R_2 be m and n respectively. Then, since R_1 and R_2 are rational, they have branch points of total order $2m-2$ and $2n-2$ respectively. Hence on the new surface R the total order of the

branch points is $2(m+n)-2$. Thus R is rational and the lemma is proved.

In the next proof we shall use the fact that R has points exterior to W. For instance, one of the two copies of γ on R crosscuts W and the other lies outside of W.

THEOREM 1. *If $f(z)$ is meromorphic in $|z|<1$, then a necessary and sufficient condition that it can be decomposed, $f(z)\equiv f_2(f_1(z))$, where $f_2(\zeta)$ is a rational function and $\zeta=f_1(z)$ is regular, bounded and univalent on $|z|<1$ is that $f(z)$ be rationally-multivalent at every point of $|z|=1$.*

Proof. We need consider only the sufficiency. If $f(z)$ is rationally multivalent along $|z|=1$ then every point of $|z|\leqq 1$ has a neighborhood (half-neighborhood for points of $|z|=1$) whose map under $w=f(z)$ can be imbedded on a rational surface. By the Heine-Borel theorem we can choose a finite number of these neighborhoods and half-neighborhoods covering $|z|<1$. If $g_2(\zeta)$ is a rational function such that $\zeta=g_2^{-1}(f(z))$ is univalent in one of these neighborhoods, then for almost all boundary points of this neighborhood $\lim g_2^{-1}(f(z))$ exists as z approaches the boundary along the interior normal. Thus the radial limit $\lim_{r\to 1} f(re^{i\theta})$ exists for almost all θ. We can therefore effect a decomposition of $|z|<1$ into a finite number of cells, bounded by Jordan arcs and Jordan cuts of $|z|<1$, such that the maps of the latter under $w=f(z)$ are Jordan arcs and Jordan cuts of the Riemann configuration W onto which $w=f(z)$ maps the circle $|z|<1$. These divide W into a finite number of simply-connected Riemann configurations (the maps of the cells of $|z|<1$) each of which can be imbedded on a rational surface. By an induction on Lemma 1, W itself can be imbedded on a rational surface R. Let $f_2(\zeta)$ be a rational function mapping the extended ζ-plane onto R. Then $f_1(z)\equiv f_2^{-1}(f(z))$ is univalent on $|z|<1$, and $f(z)\equiv f_2(f_1(z))$. Since R contains points exterior to W, the function $f_2(\zeta)$ can be chosen so that $f_1(z)$ is regular and bounded in $|z|<1$.

In each of the following two theorems one part is trivial, while the other part is an immediate consequence of Theorem 1.

THEOREM 2. *If $f(z)$ is meromorphic in $|z|<1$, and continuous in $|z|\leqq 1$, then a necessary and sufficient condition that it can be decomposed as in Theorem 1 is that $f(z)$ be power-multivalent at every point of $|z|=1$.*

THEOREM 3. *A necessary and sufficient condition that $f(z)$ be meromorphic in $|z|\leqq 1$ is that $f(z)$ be decomposable as in Theorem 1 with $f_1(z)$ regular and univalent in $|z|\leqq 1$.*

Theorem 2 follows from the fact that if $f(z)$ is continuous on $|z|\leqq 1$ then rational-multivalency along $|z|=1$ implies power-multivalency. For if the map of a half-neighborhood of z_0 ($|z_0|=1$) can be imbedded in a rational surface R, then, by the continuity of $f(z)$ at z_0, a small enough half-neighborhood will involve at most one branch point of R. Theorem 3 is even more direct,

for a function is power-multivalent at any point where it is meromorphic, and thus $f(z)$ is power-multivalent on the closed circle $|z| \leqq 1$.

If we view the decomposition of $f(z)$ with regard to the kind of plane simply-connected region for which $\zeta = f_1(z)$ is a mapping function, then Theorems 1, 2 and 3 correspond in that order to general simply-connected regions, simply-connected regions all of whose boundary points are accessible (attainable), and Jordan regions bounded by analytic curves. The corresponding theorem for general Jordan regions will be developed in §5.

THEOREM 4. *If $f(z)$ is meromorphic in $|z| < 1$, then a necessary and sufficient condition that $f(z)$ can be decomposed as in Theorem* 1 *is that there exist a univalent function $z = f_3(\zeta)$ mapping a simply-connected region G of the ζ-plane onto $|z| < 1$ such that $f(f_3(\zeta))$ is meromorphic in a simply-connected region G_1 of the ζ-plane containing the closure $\overline{G}$ of G.*

Proof. The condition is obviously necessary; we need only take $f_3(\zeta) \equiv f_1^{-1}(\zeta)$. On the other hand, suppose that there exists such a function $z = f_3(\zeta)$. Consider an analytic Jordan curve J in G_1 containing $\overline{G}$ in its interior. (We can find J by mapping G_1 onto the unit circle $|z| < 1$. The map of $\overline{G}$ is bounded away from the circumference and hence is interior to some circle $|z| = r < 1$; then J may be taken as the image of $|z| = r$.) We have $f(f_3(\zeta))$ meromorphic on and within J, and thus by Theorem 3 we can find a decomposition $f(f_3(\zeta)) \equiv f_2(g_1(\zeta))$ where f_2 and g_1 are rational and univalent respectively. Then

$$f(z) \equiv f_2(g_1(f_3^{-1}(z))) \equiv f_2(f_1(z))$$

where $f_1(z) \equiv g_1(f_3^{-1}(z))$ is univalent on $|z| < 1$.

4. **Miscellaneous decomposition theorems.** In this section we shall prove two interesting theorems which are, in a certain sense, natural predecessors of the theorems of the next section.

THEOREM 5. *If $f(z)$ is meromorphic in $|z| < 1$, and if the set S of limiting values of $f(z)$ as $|z| \to 1$ contains no value assumed by $f(z)$ in $|z| < 1$, then $f(z)$ can be written $f(z) \equiv f_2(f_1(z))$ where $\zeta = f_1(z)$ is a circular rational function, and $w = f_2(z)$ is a mapping function for a simply-connected region in the w-plane whose boundary is precisely the set S.*

Proof. Let O be the set of values assumed by $f(z)$ in $|z| < 1$. By hypothesis the boundary of O is S; moreover S is a continuum. Therefore O is a simply-connected region of the w-plane. If $\zeta = f_2^{-1}(w)$ maps O conformally onto $|\zeta| < 1$, then $\zeta = f_1(z) \equiv f_2^{-1}(f(z))$ maps $|z| < 1$ into $|\zeta| < 1$ and the set of limiting values of $\zeta = f_1(z)$ as $|z| \to 1$ lies on $|\zeta| = 1$. From this it follows that $f_1(z)$ is a circular rational function. For $f_1(z)$ can have only a finite number of zeros in $|z| < 1$, and if $g(z)$ is a circular rational function having the same set of zeros, then $\zeta = f_1(z)/g(z)$ is a regular function which does not vanish in $|z| < 1$, and whose limiting values as $|z| \to 1$ again lie on $|\zeta| = 1$. But this function

must be identically constant. Thus $\zeta = f_1(z)$ is a circular rational function, and $f(z) \equiv f_2(f_1(z))$, which completes the proof of the theorem.

Now if $f(z)$ is regular in $|z| < 1$, and if its limiting values lie on a Jordan curve of the w-plane, we do not need to assume that the limiting set contains no value taken on by $f(z)$ in $|z| < 1$. The proof of this theorem depends on the following geometric formulation of the principle of maximum modulus:

If $f(z)$ is regular in $|z| < 1$, and if the limiting values of $f(z)$ as $|z| \to 1$ form a bounded set S in the w-plane, then every value which $f(z)$ assumes in $|z| < 1$ is separated by S from infinity.

For if a value $w_0 = f(z_0)$ with $|z_0| < 1$ is not separated from infinity by S, there is an arc joining w_0 to ∞ which contains no point of S save perhaps the endpoint w_0. If w traces this arc starting from w_0, then z traces an arc lying in $|z| < 1$, which contradicts the boundedness of $f(z)$.

THEOREM 6. *If $f(z)$ is regular in $|z| < 1$, and if the set S of limiting values of $f(z)$ as $|z| \to 1$ lies on a Jordan curve Γ of the w-plane, then $f(z)$ can be expressed $f(z) \equiv f_2(f_1(z))$ where $\zeta = f_1(z)$ is a circular rational function, and $w = f_2(\zeta)$ is a mapping function for the interior of Γ.*

For all the points of Γ are accessible from infinity, and thus, by the geometric principle formulated above, the values of $f(z)$ in $|z| < 1$ lie inside Γ. If $\zeta = f_2^{-1}(w)$ maps this interior Jordan region onto $|\zeta| < 1$, we have just as in Theorem 5 that $\zeta = f_1(z) \equiv f_2^{-1}(f(z))$ is a circular rational function, and that $f(z) \equiv f_2(f_1(z))$.

COROLLARY 1. *In either of the above two theorems if some value is assumed by $f(z)$ precisely once in $|z| < 1$, or is a limiting value of $f(z)$ at precisely one point of $|z| = 1$, then $w = f(z)$ is a mapping function for a simply-connected region of the w-plane whose boundary is the set S.*

COROLLARY 2. *In either of the above two theorems $f(z)$ can be expressed as in Theorem 1 as a rational function of a mapping function.*

Corollary 1 is obvious, and Corollary 2 follows from the fact that $f(z)$ is locally univalent along $|z| = 1$ in both theorems.

5. **Decompositions involving mapping functions for Jordan regions.** In this section we shall consider the generalization of Theorem 6 in which the conditions of that theorem are assumed on only part of the circumference $|z| = 1$.

THEOREM 7. *If $f(z)$ is meromorphic in $|z| < 1$, and if the limiting values of $f(z)$ as $|z| \to 1$ along an arc γ of $|z| = 1$ lie on a Jordan arc Γ of the w-plane, then $f(z)$ is continuous on the interior of γ.*

Proof. Let z_0 be an interior point of γ and let w_0 be a limiting value of $f(z)$

at z_0. Choose a sequence $\{z_i\}$ in $|z|<1$, converging to z_0, and having the following properties: (1) $f(z_i)\to w_0$, (2) all $f(z_i)$ lie off Γ and on the same side of Γ (if w_0 is an endpoint of Γ this condition is vacuous), (3) no $f(z_i)$ is a value of $f(z)$ at a zero of $f'(z)$. All of this can be gotten from an original sequence satisfying (1) by slight displacements and the selection of a subsequence.

Let J be a Jordan arc of the w-plane having no point in common with Γ except for an endpoint at w_0, and let J contain all the $f(z_i)$ and contain no value w' assumed by $f(z)$ at a zero of $f'(z)$. (This last is possible because such values w' are at most denumerable and can accumulate only on Γ.)

Now consider a half-neighborhood R of z_0, and its bounding arc γ' on $|z-z_0|=r$. By Fatou's theorem and the theorem of F. and M. Riesz, we may take r so that at the endpoints of γ', $f(z)$ has definite radial limits, different from each other and from w_0. But then there are at most a finite number of points on γ' where $f(z)$ takes the value w_0, and we remove these by changing γ' to a new circular arc with the same endpoints, lying in R sufficiently close to γ'. We shall call this new arc and its neighborhood of z_0 respectively γ' and R.

Since $f(\gamma')$ does not contain w_0, it has no point in common with a sufficiently small circle $|w-w_0|<\rho$. All of J from some point on lies in this circle, and contains the maps $f(z_i)$ of all z_i from some i on. Let J now denote this subarc, and the z_i this subsequence. Then no point of J except w_0 is a limiting value of $f(z)$ as z approaches the boundary of R. Therefore, starting at each z_i, as $f(z)$ traces J in both directions from $f(z_i)$ the point z traces in R a simple arc J_i containing z_i. Since all the limiting values of $f(z)$ lie on Γ, there can be at most a finite number of the J_i, and some one, say J_1, contains an infinite number of the z_i. But then J_1 is a Jordan arc with endpoint z_0, because, by just the argument we have been giving, *every* half-neighborhood of z_0 contains all of J_1 from some point on.

Thus for every limiting value w_0 at z_0 there is a Jordan cut J to z_0 such that as $z\to z_0$ on J, $f(z)\to w_0$. But then all limiting values at z_0 are the same (Lindelöf) and $f(z)$ is continuous at z_0, which proves the theorem.

The following fact will prove useful in the next theorem. If w_0 is not an endpoint of Γ, let J' be a Jordan curve lying in the circle $|w-w_0|<\rho$, such that a subarc of Γ containing w_0 crosscuts the interior of J' and divides it into two Jordan regions R_1 and R_2. We assert that the map of R under $f(z)$ covers at least one of R_1 and R_2. For no limiting value of $f(z)$ along the boundary of R lies in R_1 or R_2. And R contains points (of J_1) mapping into one of R_1 and R_2.

THEOREM 8. *If $f(z)$ is meromorphic in the unit circle $|z|<1$ and continuous on a closed arc γ of the circumference $|z|=1$, and if the values which $f(z)$ assumes on γ lie on a Jordan arc Γ of the w-plane, then $f(z)$ is power-multivalent at every point of γ_1, the interior arc of γ.*

Proof. Our general scheme of proof here is to continue $f(z)$ over γ_1 in such a way that the new function $F(z)$, defined on a region containing γ_1, is an interior transformation. The power-multivalency of $f(z)$ along γ_1 is an immediate consequence of this property.

Let R_w be a Jordan region in the w-plane which contains Γ as a crosscut. Let γ' be a Jordan crosscut of $|z|<1$ having the same endpoints as γ, and let γ' be chosen so that, in the Jordan region R_z' bounded by γ and γ', $f(z)$ takes only values lying in R_w. (We can do this as follows. Since Γ can always be made a little longer, we can assume that only interior points of Γ are covered by $f(z)$ on γ_1. But then, by the continuity of $f(z)$ on γ, each point of γ_1 has a half-neighborhood whose map lies in R_w; and we can draw γ' on these half-neighborhoods.) Let R_z'' be the reflection of R_z' in $|z|=1$, and let R_z be the Jordan region $R_z'+\gamma_1+R_z''$. We now define a function $F(z)$ over R_z as follows. Let the two Jordan regions into which Γ divides R_w be R_w' and R_w'', and consider a homeomorphism between the closures of R_w' and R_w'' which leaves Γ invariant. Let the value of $F(z)$ at a point of R_z'' be defined as the homeomorph of the value of $f(z)$ at the inverse point in R_z'. And if z lies in $R_z'+\gamma_1$, let $F(z)=f(z)$.

We shall show that $w=f(z)$ is an interior transformation. First, $F(z)$ is obviously continuous. Second, $F(z)$ maps no continuum of R_z onto a point w_1 of R_w. For the set of points of $R_z'+R_z''$ which map onto w_1 is isolated in $R_z'+R_z''$ (by the definition of $F(z)$ and the analyticity of $f(z)$ in R_z'). And if $f(z)$ is not identically constant, the set of points on γ_1 which map onto w_1 is nondense on γ_1. Finally, $F(z)$ takes any open set in R_z into an open set of R_w. For the map of a neighborhood of z (in R_z) contains a neighborhood (in R_w) of $F(z)$. This follows for z in $R_z'+R_z''$ from the analyticity of $f(z)$ in R_z' and the definition of $F(z)$, and for z on γ_1 from the remark at the end of the previous theorem and the definition of $F(z)$. But then $F(z)$ has the properties which define an interior transformation; and therefore, by the fundamental theorem on interior transformations[14] $F(z)$ has at every point of R_z the property which we have called power-multivalency. In particular, $f(z)$ is power-multivalent along γ_1 (with respect to half-neighborhoods in $|z|<1$).

We have at once from Theorems 1, 7 and 8:

THEOREM 9. *If $f(z)$ is meromorphic in $|z|<1$, then a necessary and sufficient condition that $f(z)$ be decomposable as in Theorem* 1, *where $\zeta=f_1(z)$ is a mapping function for a Jordan region of the ζ-plane, is that for every z_0 on $|z|=1$ there is an integer n and an arc γ of $|z|=1$ containing z_0, such that for some w_0 the limiting values of $[f(z)-w_0]^{1/n}$ as $|z|\to 1$ along γ lie on a Jordan arc Γ of the w-plane (w_0 and Γ depending on z_0).*

It will be easily verified that Theorems 7 and 8 remain valid if we replace

---

[14] See Stollow, *Principes Topologiques de la Théorie des Fonctions Analytiques*, pp. 108–116.

the single arc Γ by a finite collection of Jordan arcs having only a finite number of points of intersection. Theorem 9 could of course be rephrased in this respect, but Theorem 9 is complicated enough already, and we shall use this generalization only in the following theorem.

THEOREM 10. *If $f(z)$ is meromorphic on $|z|<1$, and if the limiting values of $f(z)$ as $|z|\to 1$ all lie on a closed curve of the w-plane with at most a finite number of self-intersections, then $f(z)$ can be decomposed as in Theorem* 1, *where $\zeta=f_1(z)$ is a mapping function for a Jordan region.*

This theorem follows immediately from Theorem 1 and from Theorems 7 and 8 in the more general form mentioned above.

We close this section with the generalization of Theorem 5 in which the conditions of that theorem are assumed only locally along the circumference $|z|=1$. With these weakened assumptions we can still effect, as in Corollary 2 of that theorem, a decomposition of the kind in Theorem 1.

THEOREM 11. *If $f(z)$ is meromorphic in $|z|<1$, and if for every z_0 on $|z|=1$ there is a half-neighborhood R with boundary arc γ on $|z|=1$, such that the set of limiting values of $f(z)$ as $z\to\gamma$ in R contains no value assumed by $f(z)$ in R, then $f(z)$ can be decomposed as $f(z)\equiv f_2(f_1(z))$ where $\zeta=f_1(z)$ is a univalent function mapping $|z|<1$ conformally onto a bounded simply-connected region G of the ζ-plane, and $f_2(\zeta)$ is a rational function which is locally univalent (no multiple pole of $f_2(\zeta)$ or zero of $f_2'(\zeta)$) on the boundary of G.*

Proof. The half-neighborhood R is bounded by the arc γ of $|z|=1$ and a circular arc γ' of $|z-z_0|=r$. By hypothesis we may assume that the values of $f(z)$ on the interior of γ' are distinct from the limiting values of $f(z)$ along γ. The latter set forms a continuum S of the w-plane which contains no point of the set O mapping from R. Then O lies in a simply-connected region O' bounded by S, and if $w=f_2(\zeta)$ maps the unit circle $|\zeta|<1$ onto O', then the limiting values of $\zeta=f_1(z)\equiv f_2^{-1}(f(z))$ along γ lie on the Jordan curve $|\zeta|=1$. By Theorems 7 and 8, $f_1(z)$ is continuous and power-multivalent along γ, and since by hypothesis $f(z)$ assumes no value in R which it assumes on γ, $f_1(z)$ is locally univalent along the interior of γ. Hence $f(z)$ is locally univalent along γ, and likewise along the whole of $|z|=1$. But then the required decomposition is possible by Theorem 1.

6. **The function $f(z)$ regular and bounded in $|z|<1$.** In this section we shall consider the decomposition of $f(z)$ with respect to special classes of rational functions. It is evident that we can make no restriction on $f_2(\zeta)$ so long as we leave $f(z)$ meromorphic on $|z|<1$. But if we consider $f(z)$ regular and bounded in $|z|<1$, then in certain cases we can confine $f_2(\zeta)$ to polynomials or to circular rational functions, as we shall proceed to show.

LEMMA 2. *If a simply-connected Riemann configuration W can be divided by a crosscut into two Riemann configurations each of which can be imbedded in*

a surface of type II, *and if all the boundary points of W are attainable, then W itself can be so imbedded.*

Proof. This lemma is similar to Lemma 1 but more difficult. Let γ with endpoints A and B be the crosscut dividing W into W_1 and W_2, and let W_1 and W_2 be imbedded in surfaces (of type II) R_1 and R_2 respectively. We shall take R_1 and R_2 over the same circle $|w| \leq M$ of the w-plane. If we join R_1 and R_2 just as in Lemma 1, we shall have W imbedded in a disc surface R which has one too many branch points to be a surface of type II.

Let J_1 and J_2 be Jordan cuts on R_1 and R_2 respectively which run from B to the circumference $|w| = M$ and have no points in common with W_1 and W_2 respectively. Such J_1 and J_2 exist, for if a bounded plane simply-connected region G has all its boundary points attainable, then every boundary point of G can be joined to infinity along a Jordan arc having no point in common with G. In R, J_1+J_2 forms a non-dividing Jordan crosscut. If we can find a surface R' of type II which also contains J_1+J_2 as a Jordan crosscut, then by cutting each surface along J_1+J_2 and joining the left of J_1+J_2 on R to its right on R' and vice versa, we construct a new surface which contains W and which is obviously of type II.

To find such a surface R we consider again R_1 and R_2. The surface R_1 contains J_1 and R_2 contains J_2. Now if we can find a cut J which lies on each of R_1 and R_2 and runs from B to the boundary, and which touches J_1 and J_2 only at B, then by cutting R_1 and R_2 along J and joining them together around B (the left of J on R_1 to its right on R_2 and vice versa) the resulting surface will be of type II and will contain J_1+J_2 as a Jordan crosscut. We accordingly seek such a J.

Let C be a point on γ having maximum distance from B (measured on the w-plane) and let J'' be the straight line segment from C to the boundary $|w| = M$ in the direction of a line from B to C. (If J'' meets a branch point of R_1 it may be continued on any one sheet joined there.) Now if J' is the arc of γ between B and C, then the Jordan cut $J = J'+J''$ lies on both R_1 and R_2. But in general $J = J'+J''$ will intersect J_1 on R_1 at points other than B.

We avoid this difficulty by suitably enlarging R_1. If we cut R_1 along J and join it to a model of R_1 similarly cut (the left of J on R_1 to its right on the model, and vice versa) we get a surface of type II containing two images each of J_1 and J. We assert that if n models of R_1 be added to R_1 in this way, each cut along J, and all joined cyclically around B (the left of J on R_1 to its right on the first model, its left on the first model to its right on the second model, and so on until finally its left on the nth model to its right on R_1) then for some integer n the resulting surface of type II, R_1', contains a pair J_1 and J having only the point B in common.

For on R_1, J_1 does not cross J' (since the points on one side of J' lie in W_1, while J_1 is outside of W_1). Hence a subarc of J_1 between two intersections of J_1

with J'' has a definite integral order on R_1 with respect to J', and since J_1 is a Jordan arc, there is an integer N which is greater than the order of every such subarc. If N models of R_1 are added as described above, the resulting surface (of type II) has $N+1$ copies of J which are ordered cyclically so that if a given copy of J_1 meets the nth and the $(n+j)$th copies of J'', then on R_1 a subarc of J_1 has order of at least j with respect to J'. But then there is some copy of J'', and hence of $J=J'+J''$, which the given copy of J_1 does not meet.

Similarly we find a surface of type II, R_1', containing $J=J'+J''$ and J_2 with only the point B in common. If we now join R_1' and R_1' as indicated earlier, we have the required surface R', which completes the proof of the lemma.

THEOREM 12. *If $f(z)$ is regular and bounded in $|z|<1$ then Theorem 2 holds with $f_2(\zeta)$ taken either as a polynomial or as a circular rational function.*

The proof for the representation of $f(z)$ in terms of the circular rational function follows from Lemma 2 just as the proof of Theorem 1 followed from Lemma 1, and consists mainly in showing that the Riemann configuration W onto which $w=f(z)$ maps the unit circle $|z|<1$ can be imbedded in a surface of type II. But any surface of type II can be completed to a rational surface of a polynomial by adding the exteriors of the n discs $|w|\leqq M$ joined around a single branch point of order $n-1$ at infinity. Therefore the Riemann configuration W can be imbedded in a rational surface of a polynomial, and this completes the proof of the theorem.

Classified by the type of mapping function occurring in the decomposition, this theorem corresponds to Theorem 2 rather than to Theorem 1. The analogues of Theorems 3 and 10 are also theorems, and follow immediately from Theorem 12 (and Theorems 8 and 9). The analogue of Theorem 1, however (and hence also the analogue of Theorem 4), is not a theorem, as we shall show by the following example.

Consider the circles $|w\pm i|=1$ on the w-plane. If both are cut at $w=0$, and are suitably placed on the rational surface R having two sheets and branch points at $\pm i$, they can be joined to form a simple smooth closed curve which is described like a figure-eight. Now consider the circles $|w\pm i|=1\pm\frac{1}{2}$. On R these can be taken to form a second figure-eight, having no point in common with the first, such that the two together bound a band S in the shape of a figure-eight. This band is doubly-connected and is homeomorphic with the annulus $\frac{1}{2}<|z|<1$ in the z-plane. We can define a simply-connected region in the annulus by means of a spiral with an infinite number of turns which approach $|z|=1$. Consider, for instance, the spiral $z=(1-e^{-x})(\cos x+i\sin x)$, $0<x<\infty$. That part of it which lies in the annulus defines such a simply-connected region W', and this in turn defines in S a simply-connected region

W bounded by a curve which traces an infinite succession of figure-eights, and approaches the figure-eight over the pair of circles $|w \pm i| = 1$.

But then any surface R' of type II which contains W must contain a simple closed curve consisting of the figure-eight over the circles $|w \pm i| = 1$ traced a finite number of times. This is impossible, for such a curve would not divide R', whereas every simple closed curve on a surface of type II divides it.

The region W can be mapped from the unit circle $|z| < 1$ by a rational function of a mapping function. But since W cannot be imbedded in a surface of type II, the rational function cannot be taken as a polynomial or a circular rational function. This completes our demonstration.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

RESTRICTED LIE ALGEBRAS OF CHARACTERISTIC p

BY

N. JACOBSON

In an earlier paper(1) [3] we noted certain identities which connect addition, scalar multiplication, commutation ($[ab]=ab-ba$), and pth powers in an arbitrary associative algebra of characteristic p ($\neq 0$). These lead naturally to the definition of a class of abstract algebras called restricted Lie algebras which in many respects bear a closer relation to Lie algebras of characteristic 0 than ordinary Lie algebras of characteristic p.

As is shown in the present paper any restricted Lie algebra $\mathfrak{L}$ may be obtained from an associative algebra by using the operations mentioned above. In fact $\mathfrak{L}$ determines a certain associative algebra $\mathfrak{U}$, called its u-algebra, such that $\mathfrak{L}$ is isomorphic to a subalgebra of $\mathfrak{U}_l$, the restricted Lie algebra defined by $\mathfrak{L}$; and if $\mathfrak{B}$ is any associative algebra such that $\mathfrak{B}_l$ contains a subalgebra homomorphic to $\mathfrak{L}$ and $\mathfrak{B}$ is the enveloping algebra of this subset then $\mathfrak{U}$ is homomorphic to $\mathfrak{B}$. The algebra $\mathfrak{U}$ has an anti-automorphism relative to which the elements corresponding to those in $\mathfrak{L}$ are skew. For ordinary Lie algebras an algebra having these properties has been defined by G. Birkhoff [2] and by Witt [5]. In their case however, the associative algebra has an infinite basis even when the Lie algebra has a finite basis whereas here $\mathfrak{U}$ has a finite basis if and only if $\mathfrak{L}$ has. Consequently every restricted Lie algebra $\mathfrak{L}$ with a finite basis has a (1-1) representation by finite matrices. The theory of representations of $\mathfrak{L}$ can be reduced to that of the associative algebra $\mathfrak{U}$. Thus, for example, there are only a finite number of inequivalent irreducible representations.

The most natural way to obtain a restricted Lie algebra is as a derivation algebra of an arbitrary algebra $\mathfrak{A}$, i.e., as the set of transformations $D: a \to aD$ in $\mathfrak{A}$ such that

$$(a+b)D = aD + bD, \qquad (a\alpha)D = (aD)\alpha, \qquad (ab)D = (aD)b + a(bD).$$

If $\mathfrak{A}=\mathfrak{L}$ is itself restricted ($ab=[ab]$) the derivations which satisfy

$$a^pD = [[aD, \overbrace{a]a]\cdots a}^{p-1}]$$

are called restricted. They are precisely the derivations of $\mathfrak{L}$ which can be extended to derivations of the u-algebra $\mathfrak{U}$. Hence their totality is a restricted Lie algebra $\mathfrak{D}_0$. Using $\mathfrak{D}_0$ and $\mathfrak{L}$ we may define a restricted holomorph $\mathfrak{H}_0$ of $\mathfrak{L}$. $\mathfrak{H}_0$ is a restricted Lie algebra.

Presented to the Society, September 12, 1940; received by the editors June 30, 1940.

(1) Numbers in brackets refer to the bibliography at the end of the paper.

The considerations in the present paper apply for the most part to restricted Lie algebras with an infinite basis as well as to those with a finite basis. A special result, however, for Lie algebras with a finite basis is that the nilpotency of $\mathfrak{L}$ implies that of the u-algebra (§6). In a later paper we hope to discuss certain classes of simple restricted Lie algebras with a finite basis.

1. **Restricted Lie algebras. Definitions.** If $\mathfrak{A}$ is an associative algebra and the commutator $[ab]=ab-ba$, it is well known that

$$[ab] = -[ba], \qquad [a[bc]] + [b[ca]] + [c[ab]] = 0.$$

If, in addition $\mathfrak{A}$ has characteristic p ($\neq 0$) then the following identities hold:

$$(a+b)^p = a^p + b^p + s(a,b), \qquad [\overbrace{[ab]b]\cdots b}^{p}] = [ab^p],$$

where $s(a,b)=s_1(a,b)+s_2(a,b)+\cdots+s_{p-1}(a,b)$ and the $(p-i)s_i(a,b)$ is the coefficient of λ^{p-i-1} in

$$[\cdots[a,\lambda a+b],\lambda a+b],\cdots,\lambda a+b]$$

([3], and [6]). We are thus led to define a *restricted Lie algebra* $\mathfrak{L}$ over a field Φ of characteristic p as a vector space over Φ in which operations $[ab]$ and $a^{[p]}$ are defined such that

(1) $$[ab] = -[ba], \qquad [a[bc]] + [b[ca]] + [c[ab]] = 0,$$

(2) $$[a, b_1+b_2] = [ab_1] + [ab_2],$$

(3) $$[ab]\alpha = [a, b\alpha] = [a\alpha, b], \qquad \alpha \text{ in } \Phi,$$

(4) $$(a+b)^{[p]} = a^{[p]} + b^{[p]} + s(a,b),$$

(5) $$(a\alpha)^{[p]} = a^{[p]}\alpha^p,$$

(6) $$[\cdots[\overbrace{[ab]b]\cdots b}^{p}] = [ab^{[p]}].$$

A subspace $\mathfrak{B}$ of $\mathfrak{L}$ is a *subalgebra* if $\mathfrak{B}\supset b^{[p]}$ and $[b_1b_2]$ for all b, b_1, b_2 in $\mathfrak{B}$. $\mathfrak{B}$ is an *ideal* if it contains also $[ba]$ for all b in $\mathfrak{B}$, a in $\mathfrak{L}$. A correspondence $a\rightarrow a^S$ between two restricted Lie algebras is a *homomorphism* if

(7) $$(a+b)^S = a^S + b^S, \qquad (a\alpha)^S = a^S\alpha, \qquad [ab]^S = [a^Sb^S],$$
$$(a^{[p]})^S = (a^S)^{[p]}.$$

If S is (1-1) it is an *isomorphism* and if besides S is a correspondence within $\mathfrak{L}$ it is an *automorphism*.

If $x_1, x_2, \cdots$ (possibly infinite) is a basis for $\mathfrak{L}$, then $[x_ix_j]=\sum x_q\gamma_{qij}$, $x_i^{[p]}=\sum x_r\mu_{ri}$ (finite sums) where

(8) $$\gamma_{qij} = -\gamma_{qji}, \qquad \sum_q \gamma_{rqk}\gamma_{qij} + \sum_q \gamma_{rqi}\gamma_{qjk} + \sum_q \gamma_{rqj}\gamma_{qki} = 0,$$

(9) $$\sum_r \gamma_{sir}\mu_{rj} = \sum_q \gamma_{sq_{p-1}j}\gamma_{q_{p-1}q_{p-2}i}\cdots\gamma_{q_1ij}.$$

These equations are equivalent to

$$[x_ix_j] = -[x_jx_i],$$

$$[[x_ix_j]x_k] + [[x_jx_k]x_i] + [[x_kx_i]x_j] = 0, \quad [x_ix_j^{[p]}] = [\cdots [x_ix_j]\cdots x_j],$$

respectively. The γ's and μ's are the *constants of multiplication* of $\mathfrak{L}$. If $a \to a^S$ is an isomorphism between $\mathfrak{L}$ and $\mathfrak{L}^S$, $x_1^S, x_2^S, \cdots$ form a basis for $\mathfrak{L}^S$ and the x_i^S have the same constants of multiplication as the x_i's. On the other hand if $\mathfrak{M}$ is any restricted Lie algebra with basis $y_1, y_2, \cdots$ in (1-1) correspondence with the x_i such that $[y_iy_j] = \sum y_q\gamma_{qij}$, $y_i^{[p]} = \sum y_r\mu_{ri}$ then it is readily seen that the correspondence $\sum x_i\alpha_i \to \sum y_i\alpha_i$ is an isomorphism.

We have noted above that any associative algebra $\mathfrak{A}$ of characteristic p becomes a restricted Lie algebra $\mathfrak{A}_l$ when $[ab]$ is defined as $ab-ba$ and $a^{[p]} = a^p$. A homomorphism between $\mathfrak{L}$ and a subalgebra of Φ_{nl}, Φ_n the algebra of $n \times n$ matrices, is called a *representation*. Irreducibility, decomposability, equivalence, etc., of representations are defined as usual. As is well known these depend on the irreducibility, etc., of the enveloping algebra of the representing matrices.

2. **The u-algebra of a restricted Lie algebra.** Let $x_1, x_2, \cdots$ (possibly infinite) be a basis over Φ of a vector space $\mathfrak{L}$. We set $[x_ix_j] = \sum x_q\gamma_{qij}$ and $x_i^{[p]} = \sum x_r\mu_{ri}$ where the γ's are μ's satisfy (8) and (9). Then for $a = \sum x_i\alpha_i$, $b = \sum x_i\beta_i$ (finite sums) we set $[ab] = \sum x_k\gamma_{kij}\alpha_i\beta_j$. Evidently (8) implies that $\mathfrak{L}$ is a Lie algebra relative to $[ab]$. Equation (9) is equivalent to

$$[x_ix_j^{[p]}] = [\cdots [x_i\overbrace{x_j]\cdots x_j}^{p}].$$

We shall show that $\mathfrak{L}$ is a restricted Lie algebra relative to a suitable definition of $a^{[p]}$.

Let $\mathfrak{A}$ be the vector space with the basis $x_1^{\kappa_1}x_2^{\kappa_2}\cdots x_n^{\kappa_n}$, $\kappa_i \geqq 0$ integers and at least one $\kappa_i > 0$, $n = 1, 2, \cdots$. If only a finite number of monomials are being considered we may write them in terms of the same x's. A product

$$(x_1^{\varepsilon_1}x_2^{\varepsilon_2}\cdots x_n^{\varepsilon_n})(x_1^{\lambda_1}x_2^{\lambda_2}\cdots x_n^{\lambda_n})$$

is defined by repeated "straightenings," i.e., substitutions for x_ix_j when $i > j$ of the expression $x_jx_i + \sum x_k\gamma_{kij}$. It has been shown by G. Birkhoff [2] and by Witt [5] that this product is uniquely defined in $\mathfrak{A}$ and is associative.

Let $\mathfrak{B}$ be the ideal in $\mathfrak{A}$ having the basis $y_i = x_i^p - x_i^{[p]}$. Since

$$[bx_i^p] = [\cdots [bx_i]x_i]\cdots x_i] = [bx_i^{[p]}], \tag{10}$$

y_i commutes with every linear b and hence with every element of $\mathfrak{A}$.

If the term $x_1^{\varepsilon_1}\cdots x_n^{\varepsilon_n}$ has degree $\geqq p$ in x_i we may replace x_i^p by $y_i + x_i^{[p]}$. After a finite number of such substitutions we may write any $a = \sum x_1^{\varepsilon_1}\cdots x_n^{\varepsilon_n}$

$\cdot\rho_{\epsilon_1}\cdots_{\epsilon_n}$ in the form $\sum x_1^{\lambda_1}\cdots x_n^{\lambda_n}u_{\lambda_1}\cdots_{\lambda_n}$ where $\lambda_i<p$ and the u's are polynomials in $y_1, y_2, \cdots, y_n$. Thus any $b=\sum a_iy_i+\sum y_i\alpha_i$ in $\mathfrak{B}$ has the form $\sum x_1^{\lambda_1}\cdots x_n^{\lambda_n}\cdot v_{\lambda_1}\cdots_{\lambda_n}$, v a polynomial in $y_1, y_2, \cdots, y_n$ with no constant term. Now

$$x_1^{\lambda_1}\cdots x_n^{\lambda_n}y_1^{m_1}\cdots y_n^{m_n} = x_1^{\lambda_1}(x_1^p - x_1^{[p]})^{m_1}\cdots x_n^{\lambda_n}(x_1^p - x_1^{[p]})^{m_n}$$
$$= x_1^{\lambda_1+pm_1}\cdots x_n^{\lambda_n+pm_n} + \cdots$$

where the terms not indicated have degree $<\sum\lambda_i+p\sum m_i$. Consider the terms of maximum degree $N=\sum\lambda_i+p\sum m_i$ in b where we suppose $b\neq 0$ and hence one of the v's is $\neq 0$. Since at least one $m>0$, $N>p$. Two terms of maximum degree are different if $(\lambda_1, \cdots, \lambda_n; m_1, \cdots, m_n)\neq(\lambda_1', \cdots, \lambda_n'; m_1', \cdots, m_n')$. Hence these terms occur only once and can not cancel off. It follows that when b is written in its normal form $\sum x_1^{\nu_1}\cdots x_n^{\nu_n}\rho_{\epsilon_1}\cdots_{\epsilon_n}$ at least one of the x's has degree $>p$. Thus the classes $\{x_1^{\lambda_1}\cdots x_n^{\lambda_n}\}$, $\lambda_i<p$, omitting $\{x_1^0\cdots x_n^0\}$, determined by the elements $x_1^{\lambda_1}\cdots x_n^{\lambda_n}$ modulo $\mathfrak{B}$ form a basis for the difference algebra $\mathfrak{U}=\mathfrak{A}/\mathfrak{B}$.

Since the classes $\{x_1\}, \{x_2\}, \cdots$ are linearly independent and $[\{x_i\}\{x_j\}] = \sum\{x_q\}\gamma_{qij}$ the correspondence $\sum x_i\alpha_i\to\sum\{x_i\}\alpha_i$ is an isomorphism between $\mathfrak{L}$ and an ordinary Lie subalgebra $\{\mathfrak{L}\}$ of $\mathfrak{U}_l$. Since $\{x_i\}^p=\{x_i^p\}=\{x_i^{[p]}\} = \sum\{x_r\}\mu_{ri}$, $\{\mathfrak{L}\}$ is a (restricted) subalgebra of $\mathfrak{U}_l$. Hence if we define $(\sum x_i\alpha_i)^{[p]}$ as the element corresponding to $\{\sum x_i\alpha_i\}^p$, $\mathfrak{L}$ becomes a restricted Lie algebra in which $x_i^{[p]}$ is as originally given.

If $\mathfrak{L}$ is a restricted Lie algebra to begin with, then the correspondence $\sum x_i\alpha_i\to\{\sum x_i\alpha_i\}$ is an isomorphism. Now suppose $\sum x_i\alpha_i\to\sum\bar{x}_i\alpha_i$ is a homomorphism between $\mathfrak{L}$ and $\bar{\mathfrak{L}}$, a subalgebra of $\mathfrak{B}_l$, where we suppose that any element of $\mathfrak{B}$ is a polynomial in the elements of $\bar{\mathfrak{L}}$. We have

$$\bar{x}_i\bar{x}_j = \bar{x}_j\bar{x}_i + \sum \bar{x}_q\gamma_{qij}, \qquad \bar{x}_i^p = \sum \bar{x}_r\mu_{ri}.$$

The first set of equations implies that $\sum x_1^{\nu_1}\cdots x_n^{\nu_n}\rho_{\epsilon_1}\cdots_{\epsilon_n}\to\sum\bar{x}_1^{\nu_1}\cdots\bar{x}_n^{\nu_n}\rho_{\epsilon_1}\cdots_{\epsilon_n}$ is a homomorphism between $\mathfrak{A}$ and $\mathfrak{B}$(2). Because of the second set of equations $x_i^p-x_i^{[p]}$ are mapped into 0 and so our correspondence induces a homomorphism between $\mathfrak{U}=\mathfrak{A}/\mathfrak{B}$ and $\mathfrak{B}$. This mapping is an extension of the homomorphism between $\mathfrak{L}$ and $\bar{\mathfrak{L}}$. Thus we have proved the following theorem.

THEOREM 1. *If $\mathfrak{L}$ is a restricted Lie algebra there exists an associative algebra $\mathfrak{U}$ having the following properties*: 1. *$\mathfrak{L}$ is isomorphic to a subalgebra $\{\mathfrak{L}\}$ of $\mathfrak{U}_l$.* 2. *$\mathfrak{U}$ is the enveloping algebra of $\{\mathfrak{L}\}$.* 3. *If $\mathfrak{L}$ is homomorphic to a subalgebra $\bar{\mathfrak{L}}$ of any $\mathfrak{B}_l$ where $\mathfrak{B}$ is the enveloping algebra of $\mathfrak{L}$, then $\mathfrak{U}$ is homomorphic to $\mathfrak{B}$.*

This theorem shows that conditions (1) to (6) are characteristic of the functions $a+b$, $a\alpha$, $[ab]$ and a^p in an associative algebra of characteristic p. The above considerations show also that equations (8) and (9) on the con-

(2) See [2] or [5].

stants γ and μ insure that the vector space $\mathfrak{L}$ be a restricted Lie algebra. If P is an extension of the field Φ the extended vector space of elements of the form $\sum x_i\xi_i$, ξ in P, is a Lie algebra over P. We denote this extended algebra as $\mathfrak{L}P$ since, as is easily shown, it does not depend on the particular choice of basis. In the remainder of the paper we shall denote $a^{(p)}$ by a^p when there is no risk of confusion and shall call $\mathfrak{U}$ the u-algebra of $\mathfrak{L}$.

Suppose $x_{i_1}\cdots x_{i_m}$ is a monomial in $\mathfrak{A}$ and the first x_1 in this product occurs in the r_1th place. Then we may straighten this term by interchanging x_1 successively with the r_1-1 terms in front of it and obtain a polynomial in the x's having one term $x_1x_{j_2}\cdots x_{j_m}$ of degree m. We define the rank of $x_{i_1}\cdots x_{i_m}$ inductively as (r_1-1) plus rank of $x_{j_2}\cdots x_{j_m}$. A monomial of rank 0 is said to be in canonical form. Then $i_1\geqq i_2\geqq\cdots\geqq i_m$. Consider the correspondence $a=\sum x_1^{\epsilon_1}\cdots x_n^{\epsilon_n}\rho_{\epsilon_1\cdots\epsilon_n}\to a^J=\sum(-1)^{\epsilon_1+\cdots+\epsilon_n}x_n^{\epsilon_n}\cdots x_1^{\epsilon_1}\rho_{\epsilon_1\cdots\epsilon_n}$. Evidently J is linear. We wish to show that it is an anti-automorphism. For this purpose it suffices to prove that

$$(x_{i_1}\cdots x_{i_m})^J=(-1)^m x_{i_m}\cdots x_{i_1}.$$

Suppose this holds for all products of $(m-1)$ or less x's and also for products of m x's whose ranks are less than those of the given monomial. Then

$$(x_{i_1}\cdots x_{i_m})=(x_{i_1}\cdots x_{i_{j+1}}x_{i_j}\cdots x_{i_m})+(x_{i_1}\cdots[x_{i_j}x_{i_{j+1}}]\cdots x_{i_m}),$$

where we may suppose that if the rank r of the original term is >0 that of $x_{i_1}\cdots x_{i_{j+1}}x_{i_j}\cdots x_{i_m}$ is $r-1$. We have

$$\begin{aligned}(x_{i_1}\cdots x_{i_m})^J&=(x_{i_1}\cdots x_{i_{j+1}}x_{i_j}\cdots x_{i_m})^J+(x_{i_1}\cdots[x_{i_j}x_{i_{j+1}}]\cdots x_{i_m})^J\\&=(-1)^m(x_{i_m}\cdots x_{i_j}x_{i_{j+1}}\cdots x_{i_1})\\&\quad+(-1)^{m-1}(x_{i_m}\cdots[x_{i_j}x_{i_{j+1}}]\cdots x_{i_1})\\&=(-1)^m(x_{i_m}\cdots x_{i_1}).\end{aligned}$$

Since $(x_i^{[p]}-x_i^p)^J=-(x_i^{[p]}-x_i^p)$, J sends the ideal $\mathfrak{B}$ into itself and therefore induces an anti-automorphism in $\mathfrak{U}=\mathfrak{A}/\mathfrak{B}$. The elements $\sum\{x_i\}\alpha_i$ of $\{\mathfrak{L}\}$ are skew relative to the anti-automorphism.

By property (3) of $\mathfrak{U}$ any representation of $\mathfrak{L}$ determines a representation of $\mathfrak{U}$ and conversely. Questions of irreducibility, equivalence, etc., for $\mathfrak{L}$ are reducible to the corresponding questions for $\mathfrak{U}$. If $\mathfrak{L}$ has a finite basis x_1, $x_2,\cdots,x_n$, $\mathfrak{U}$ has the basis $\{x_1^{\lambda_1}\cdots x_n^{\lambda_n}\}$ $(\lambda_i<p)$ of p^n-1 elements. Since $\mathfrak{U}$ has a (1-1) representation in some Φ_m, $m\leqq p^n$, the same is true for $\mathfrak{L}$.

THEOREM 2. *Every restricted Lie algebra with a finite basis has a* (1-1) *representation.*

The number of inequivalent irreducible representations ($\neq 0$) of $\mathfrak{U}$ is equal to the number of simple components of $\mathfrak{U}/\mathfrak{N}$, $\mathfrak{N}$ the radical of $\mathfrak{U}$. This implies

THEOREM 3. *There are only a finite number of inequivalent irreducible representations of a restricted Lie algebra with a finite basis.*

Example. $\mathfrak{L}$, the restricted Lie algebra with the basis x, y, z such that

$$[xy] = z, \qquad [xz] = [yz] = 0, \qquad x^p = y^p = z^p = z.$$

$\mathfrak{U}$ has the basis $x^i y^j z^k$, $i, j, k < p$ such that the above relations hold, it being understood that $[xy] = xy - yx$, etc. It is readily proved that $\mathfrak{U}$ is a direct sum of p algebras with bases $x^i y^j$ such that

$$xy - yx = \zeta, \qquad x^p = y^p = \zeta,$$

where $\zeta = 0, 1, \cdots, p-1$ in turn. If $\zeta = 0$ this algebra is nilpotent. Otherwise it is isomorphic to Φ_p. Hence there are $(p-1)$ inequivalent irreducible representations of $\mathfrak{L}$.

Theorem 2 is valid for Lie algebras of characteristic 0 though its proof given by Ado [1] is considerably more complicated than the present one. Theorem 3 is not true for algebras of characteristic 0. It is not known whether either of these results holds for ordinary Lie algebras of characteristic p.

3. **Ideals. Nilpotency.** If $\mathfrak{B}$ is an ideal in $\mathfrak{L}$ we define the sum, scalar product and commutator of the classes $\bar{a}$ modulo $\mathfrak{B}$ as usual by

$$\bar{a}_1 + \bar{a}_2 = \overline{a_1 + a_2}, \qquad \bar{a}\alpha = \overline{a\alpha}, \qquad [\bar{a}_1 \bar{a}_2] = \overline{[a_1 a_2]}.$$

If $a_1 - a_2 = b \in \mathfrak{B}$ then $a_1^p = a_2^p + b^p + s(a_2, b)$. Since b^p and $s(a_2, b) \in \mathfrak{B}$ we see that $a_1 \equiv a_2 (\mathfrak{B})$ implies $a_1^p \equiv a_2^p (\mathfrak{B})$. Hence the definition $(\bar{a})^p = \overline{a^p}$ is unambiguous and together with the above operations it defines $\bar{\mathfrak{L}} = \mathfrak{L}/\mathfrak{B}$, the *difference algebra* of $\mathfrak{L}$ relative to $\mathfrak{B}$, as a restricted Lie algebra. The correspondence $a \to \bar{a}$ is a homomorphism between $\mathfrak{L}$ and $\bar{\mathfrak{L}}$. Conversely we may show in the usual manner that if $\mathfrak{L}$ is homomorphic to the restricted Lie algebra $\bar{\mathfrak{L}}$, $\bar{\mathfrak{L}} = \mathfrak{L}/\mathfrak{B}$ where $\mathfrak{B}$ is the set of elements mapped into 0 by the homomorphism.

If $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are subspaces of $\mathfrak{L}$ we denote their sum as $\mathfrak{B}_1 + \mathfrak{B}_2$ and their commutator, i.e., the smallest space containing all $[b_1 b_2]$, $b_1 \in \mathfrak{B}_1$, $b_2 \in \mathfrak{B}_2$ by $[\mathfrak{B}_1 \mathfrak{B}_2]$. Conditions (1), (2) and (3) imply

$$(11) \qquad [\mathfrak{B}_1 \mathfrak{B}_2] = [\mathfrak{B}_2 \mathfrak{B}_1], \qquad [\mathfrak{B}_1 [\mathfrak{B}_2 \mathfrak{B}_3]] \subseteq [\mathfrak{B}_2 [\mathfrak{B}_3 \mathfrak{B}_1]] + [\mathfrak{B}_3 [\mathfrak{B}_1 \mathfrak{B}_2]].$$

Set $\mathfrak{L}^{[2]} = [\mathfrak{L}\mathfrak{L}], \cdots, \mathfrak{L}^{[i]} = [\mathfrak{L}^{[i-1]}\mathfrak{L}]$ and define $\mathfrak{L}^{p^k}$ to be the smallest subspace containing all a^{p^k} where $a^{p^k} = (a^{p^{k-1}})^p$. Thus $\mathfrak{L} \supseteq \mathfrak{L}^{[2]} \supseteq \mathfrak{L}^{[3]} \supseteq \cdots$ and $\mathfrak{L} \supseteq \mathfrak{L}^p \supseteq \mathfrak{L}^{p^2} \supseteq \cdots$, $\mathfrak{L}^{p^2} \subseteq (\mathfrak{L}^p)^p$, etc. Hence if we define

$$\mathfrak{L}_i = \mathfrak{L}^{[i]} + (\mathfrak{L}^{[i-1]})^p + (\mathfrak{L}^{[i-2]})^{p^2} + \cdots + \mathfrak{L}^{p^{i-1}}$$

we have $\mathfrak{L} = \mathfrak{L}_1 \supseteq \mathfrak{L}_2 \supseteq \mathfrak{L}_3 \supseteq \cdots$. By induction on j one readily establishes

$$[\mathfrak{L}^{[i]} \mathfrak{L}^{[j]}] \subseteq \mathfrak{L}^{[i+j]}.$$

Since

$$[b_1^{p^k}, b_2^{p^l}] = [\cdots [b_1^{p^k}\overbrace{b_2]b_2]\cdots b_2}^{p^l}]$$

$$= [[\cdots [\overbrace{b_1[b_1\cdots [b_1b_2]\cdots]}^{p^k}, \overbrace{b_2]b_2]\cdots b_2}^{p^l-1}]$$

we have

$$[\mathfrak{L}^{[i-k]p^k}\mathfrak{L}^{[j-l]p^l}] \leqq \mathfrak{L}^{[(i-k)p^k+(j-l)p^l]} \leqq \mathfrak{L}^{[i+j]}. \tag{12}$$

This leads readily to

$$[\mathfrak{L}_i\mathfrak{L}_j] \leqq \mathfrak{L}_{i+j}, \qquad \mathfrak{L}_i^p \leqq \mathfrak{L}_{i+1}. \tag{13}$$

In particular $\mathfrak{L}_i$ is an ideal in $\mathfrak{L}$ and $\mathfrak{L}_i \geqq [\mathfrak{L}_{i-1}\mathfrak{L}]+\mathfrak{L}_{i-1}^p$. On the other hand from the definition of $\mathfrak{L}_i$ we have $\mathfrak{L}_i \leqq [\mathfrak{L}_{i-1}\mathfrak{L}]+\mathfrak{L}_{i-1}^p$. Hence $\mathfrak{L}_i = [\mathfrak{L}_{i-1}\mathfrak{L}]+\mathfrak{L}_{i-1}^p$. It follows that if $\mathfrak{L}_{i-1}=\mathfrak{L}_i$, $\mathfrak{L}_{i-1}=\mathfrak{L}_i=\mathfrak{L}_{i+1}=\cdots$. If $\mathfrak{L}_N=0$ for N sufficiently large, $\mathfrak{L}$ is *nilpotent*. The smallest N for which this holds is called the *index* of nilpotency. If $\mathfrak{L}_N=0$, $\mathfrak{L}^{p^{N-1}}=0$ and $\mathfrak{L}^{[N]}=0$. Conversely if $\mathfrak{L}^{[r]}=0$ and $\mathfrak{L}^{p^s}=0$ it is readily seen that $\mathfrak{L}_t=0$ for $t=r+s-1$.

4. **Restricted derivations.** The most natural instances of restricted Lie algebras are the derivation algebras(3). We recall that if $\mathfrak{A}$ is an arbitrary algebra (not necessarily associative) then a derivation D is defined to be a transformation $a\to aD$ in $\mathfrak{A}$ such that

$$(a+b)D = aD+bD, \quad (a\alpha)D = (aD)\alpha, \quad (ab)D = (aD)b + a(bD). \tag{14}$$

If $\mathfrak{A}$ has characteristic p the set $\mathfrak{D}$ of these transformations is closed under addition, scalar multiplication, commutation and pth powers. Thus $\mathfrak{D}$ is a restricted Lie algebra. If $\mathfrak{A}$ is associative

$$a^pD = (aD)a^{p-1} + a(aD)a^{p-2} + \cdots + a^{p-1}(aD) = [\cdots [aD, \overbrace{a]a]\cdots a}^{p-1}](4).$$

It is therefore natural to confine our attention in the case that $\mathfrak{A}=\mathfrak{L}$ is a restricted Lie algebra to the derivations called *restricted* such that

$$a^pD = [\cdots [aD\overbrace{a]a]\cdots a}^{p-1}]. \tag{15}$$

Suppose D is a linear transformation in $\mathfrak{L}$ such that $[x_ix_j]D=[x_iD, x_j]+[x_i, x_jD]$, $x_1, x_2, \cdots$ a basis for $\mathfrak{L}$. Then D is a derivation. If $x_1^{\kappa_1}\cdots x_n^{\kappa_n}$, $\kappa_i=1, 2, \cdots$, is a basis for the Birkhoff-Witt algebra $\mathfrak{A}$ we define the linear transformation D in $\mathfrak{A}$ by setting

(3) Cf. Jacobson [3].

(4) In general

$$ba^{p-1} + aba^{p-2} + \cdots + a^{p-1}b = [\cdots \overbrace{[ba]a]\cdots a}^{p-1}].$$

See [3, p. 209].

$$(x_1^{s_1}\cdots x_n^{s_n})D = (x_1D)x_1^{s_1-1}\cdots x_n^{s_n} + \cdots + x_1^{s_1-1}(x_1D)\cdots x_n^{s_n} + \cdots + x_1^{s}\cdots x_n^{s_n-1}(x_nD). \tag{16}$$

By an induction similar to that of §2 we can show that

$$(x_{i_1}\cdots x_{i_m})D = (x_{i_1}D)x_{i_2}\cdots x_{i_m} + \cdots + x_{i_1}\cdots(x_{i_m}D), \tag{17}$$

holds. Hence D is a derivation.

Now suppose

$$x_i^{[p]}D = [\cdots[x_iD, \overbrace{x_i]\cdots x_i}^{p-1}].$$

Then $(x_i^{[p]}-x_i^p)D=0$ and D maps the ideal $\mathfrak{B}$ whose basis is $x_i^{[p]}-x_i^p$ into itself. It follows that D induces a derivation in $\mathfrak{U}=\mathfrak{A}/\mathfrak{B}$ and hence is a restricted derivation in $\mathfrak{L}$. We have shown also that any restricted derivation is determined by a derivation of the u-algebra. The converse of this is clear. Hence we have proved

THEOREM 4. *A linear transformation D in $\mathfrak{L}$ is a restricted derivation if and only if $[x_ix_j]D=[x_iD, x_j]+[x_i, x_jD]$ and $x_i^pD=[\cdots[x_iD, x_i]\cdots x_i]$ for any basis $x_1, x_2, \cdots$. Every restricted derivation is induced by a derivation of the associative u-algebra $\mathfrak{U}$ of $\mathfrak{L}$. The set of restricted derivations forms a restricted Lie algebra.*

The last statement follows immediately from the second. The set of restricted derivations will be denoted by $\mathfrak{D}_0$. A consequence of the first part of Theorem 4 is that the restricted derivation algebra of $\mathfrak{L}_P$ is $\mathfrak{D}_{0P}$.

For any three elements a, b, l in $\mathfrak{L}$ we have

$$[a+b, l] = [al] + [bl],\quad [a\alpha, l] = [al]\alpha,\quad [[ab]l] = [[al]b] + [a[bl]],$$

$$[a^pl] = [\cdots[al]\overbrace{a]a]\cdots a}^{p-1}].$$

Thus the transformations $L: a\to[al]$ are restricted derivations which we call *inner*. The other parts of (1) to (6) show that the derivations corresponding to l_1+l_2, $l\alpha$, $[l_1l_2]$, l^p are respectively L_1+L_2, $L\alpha$, $[L_1L_2]$, L^p. Hence the inner derivations form a subalgebra $\mathfrak{F}$ of $\mathfrak{D}_0$ and $\mathfrak{L}$ is homomorphic to $\mathfrak{F}$ under the correspondence $l\to L$. The elements c mapped into 0 are those which satisfy $[ac]=0$ for all a and form the *center* $\mathfrak{C}$ of $\mathfrak{L}$. Hence $\mathfrak{F}\cong\mathfrak{L}/\mathfrak{C}$. Since

$$[al]D - [aD, l] = [a, lD],$$

$[LD]\in\mathfrak{F}$ for every L in $\mathfrak{F}$ and D in $\mathfrak{D}_0$, i.e., $\mathfrak{F}$ is an ideal.

Let $\mathfrak{H}_0$ be the vector space which is a direct sum of $\mathfrak{L}$ and $\mathfrak{D}_0$. The elements U of $\mathfrak{H}_0$ are uniquely representable in the form $a+D$, a in $\mathfrak{L}$, D in $\mathfrak{D}_0$. Hence if $x_1, x_2, \cdots$ is a basis for $\mathfrak{L}$ and $D_1, D_2, \cdots$ one for $\mathfrak{D}_0$, $x_1, x_2, \cdots; D_1, D_2, \cdots$ is a basis for $\mathfrak{H}_0$.

We define commutation in $\mathfrak{H}_0$ by

$$[a + D, b + E] = [ab] + aE - bD + [DE]. \tag{18}$$

It is readily seen that this satisfies conditions (1), (2), (3)(5). We also have

$$[\cdots [\overbrace{ax_i}^{p}\cdots x_i] = [ax_i^p], \qquad [\cdots [\overbrace{Dx_i}^{p}\cdots x_i] = [Dx_i^p],$$

$$[\cdots [aD_j]\cdots D_j] = [aD_j^p], \qquad [\cdots [DD_j]\cdots D_j] = [DD_j^p],$$

or

$$[\cdots [Ux_i]\cdots x_i] = [Ux_i^p], \qquad [\cdots [UD_j]\cdots D_j] = [UD_j^p],$$

for all U in $\mathfrak{H}_0$. Hence, by §2, the definition

$$(\sum x_i\alpha_i + \sum D_j\beta_j)^p = \sum x_i^p\alpha_i^p + \sum D_j^p\beta_j^p + s(x, D) \tag{19}$$

turns $\mathfrak{H}_0$ into a restricted Lie algebra called the *restricted holomorph* of $\mathfrak{L}$. Thus $(a+D)^p = a^p + D^p + s(a, D)$ in $\mathfrak{H}_0$.

5. **Relations to ordinary Lie algebras.** Suppose $\mathfrak{L}_1$ and $\mathfrak{L}_2$ are restricted Lie algebras and $a \to a^S$ is a mapping of $\mathfrak{L}_1$ into $\mathfrak{L}_2$ such that

$$(a + b)^S = a^S + b^S, \qquad (a\alpha)^S = a^S\alpha, \qquad [ab]^S = [a^S b^S].$$

Then

$$[\cdots [xa]\cdots a]^S = [\cdots [x^S a^S]\cdots a^S] = [x^S(a^S)^p],$$
$$[x, a^p]^S = [x^S(a^p)^S].$$

Hence $(a^p)^S - (a^S)^p \in \mathfrak{C}_2$ the center of $\mathfrak{L}_2$. If $\mathfrak{C}_2 = 0$, S is a homomorphism. Next we suppose D is a derivation in $\mathfrak{L} = \mathfrak{L}_1$. This implies that

$$\begin{aligned}[\cdots [xa]\cdots a]D &= [\cdots [xDa]\cdots a] + [\cdots [[x, aD]a]\cdots a] \\ &\quad + \cdots + [\cdots [[xa]a]\cdots aD] \\ &= [xD, a^p] + [x, [\cdots [aD, a]\cdots a]],\end{aligned}$$

since if A is the transformation $x \to [xa]$ and B is the transformation $x \to [x, aD]$,

$$BA^{p-1} + ABA^{p-2} + \cdots + A^{p-1}B = [\cdots [[BA]A]\cdots A](6).$$

On the other hand $[xa^p]D = [xD, a^p] + [x, a^pD]$. Hence $a^pD - [\cdots [aD, a]\cdots a]$ is in the center $\mathfrak{C}$ of $\mathfrak{L}$. If $\mathfrak{C} = 0$ every derivation is restricted.

If $\mathfrak{B}$ is a subset of $\mathfrak{L}$ closed with respect to addition, scalar multiplication and commutation, then $\mathfrak{B}^* = \mathfrak{B} + \mathfrak{B}^p + \mathfrak{B}^{p^2} + \cdots$ is the smallest subalgebra of $\mathfrak{L}$ containing $\mathfrak{B}$. If $[\mathfrak{B}\mathfrak{L}] \leq \mathfrak{B}$, $\mathfrak{B}^*$ is an ideal.

(5) Cf. Zassenhaus [6, p. 57].

(6) See Footnote 4.

An ordinary Lie algebra $\mathfrak{A}$ whose center is 0 may always be imbedded in a restricted Lie algebra. For let $\mathfrak{D}$ be the derivation algebra of $\mathfrak{A}$. Since the center of $\mathfrak{A}$ is 0 the set $\mathfrak{F}$ of inner derivations forms an ideal of $\mathfrak{D}$ (regarded as an ordinary Lie algebra) isomorphic to $\mathfrak{A}$. If $\mathfrak{F}^*=\mathfrak{F}+\mathfrak{F}^p+\cdots$, $\mathfrak{F}^*$ is a restricted Lie algebra containing $\mathfrak{F}$. Since $[u^{p^i}v^{p^j}]=[\cdots[[u[\cdots[uv]\cdots],v]\cdots v]\in\mathfrak{F}$ for any u, v in $\mathfrak{F}$ the ordinary Lie algebra $\mathfrak{F}^*/\mathfrak{F}$ is commutative.

Suppose $\mathfrak{L}$ is a restricted Lie algebra and $a\to A$ an absolutely irreducible representation of the ordinary Lie algebra determined by $\mathfrak{L}$, i.e., $a\alpha\to A\alpha$, $a+b\to A+B$, $[ab]\to[AB]$. We assume also that the representing matrices all have trace 0 and $p\nmid m$ where $m\times m$ are the dimensions of the matrices. We assert that our representation is one of the restricted Lie algebra. For if $a^p\to B$ we have $[XB]=[\cdots[XA]\cdots A]=[XA^p]$. Hence $A^p-B=1\rho$ and since tr $A^p=$tr $B=0$, $A^p=B$.

6. **Algebras with a finite basis.** In this section we suppose $\mathfrak{L}$ has a finite basis. For any element a there is a least integer m such that $a, a^p, \cdots, a^{p^{m-1}}$ are linearly independent but a^{p^m} depends on $a, \cdots, a^{p^{m-1}}$. Then $a^{p^m}+a^{p^{m-1}}\alpha_1+\cdots+a\alpha_m=0$ or $f(a)=0$ where $f(\lambda)=\lambda^{p^m}+\lambda^{p^{m-1}}\alpha_1+\cdots+\lambda\alpha_m$. It follows that $a^{p^{m+1}}=(a^{p^m})^p$, $a^{p^{m+1}}, \cdots$ are linear combinations of $a, \cdots, a^{p^{m-1}}$ and hence these elements form a basis for the subalgebra generated by a.

A polynomial having the form of $f(\lambda)$ has been called a p-polynomial by Ore. The following facts were established by Ore: (1) A necessary and sufficient condition that $f(\lambda)$ be a p-polynomial is that its roots form a modulus (group under addition) with each root having multiplicity p^k, k fixed. (2) Any polynomial $\phi(\lambda)$ is a factor of a p-polynomial $f(\lambda)$. The $f(\lambda)$ of least degree with leading coefficient 1 is unique and is a divisor of any other p polynomial divisible by $\phi(\lambda)$(7). Thus suppose $a\to A$ is a (1-1) representation of $\mathfrak{L}$ and $\phi(\lambda)$ is the minimum polynomial of the linear transformation (or matrix) A. Then the p-polynomial $f(\lambda)$ associated with a is the one of least degree divisible by $\phi(\lambda)$.

An element a is *nilpotent* if $a^{p^m}=0$ for some p^m. The least integer p^m for which this holds is the *index* of a.

THEOREM 5. *If $\mathfrak{L}$ is a Lie algebra with a finite basis and contains only nilpotent elements, then $\mathfrak{L}$ is nilpotent.*

Since $\mathfrak{L}$ has a finite basis the index of any a is $\leq p^n$ where n is the dimensionality of $\mathfrak{L}$. Hence $a^{p^n}=0$ and if A is the linear transformation $x\to[xa]$, $A^{p^n}=0$. Thus $\mathfrak{L}^{p^n}=0$ and as has been shown by Zorn [8], $\mathfrak{L}$ is nilpotent when regarded as an ordinary Lie algebra, i.e., $\mathfrak{L}^{[m]}=0$. It follows as in §4 that $\mathfrak{L}$ is nilpotent.

The algebra $\mathfrak{L}>\mathfrak{L}_2=\mathfrak{L}^{[2]}+\mathfrak{L}^p$. If $\mathfrak{M}$ is a subspace such that $\mathfrak{L}\geqq\mathfrak{M}\geqq\mathfrak{L}_2$, $\mathfrak{M}$ is a nilpotent ideal and $\mathfrak{L}/\mathfrak{M}$ is commutative and has all of its elements $\neq 0$ nil-

(7) Ore [4, p. 581].

potent of index p. We choose $\mathfrak{M}$ so that dim $\mathfrak{M}=n-1$ and let $x_2, x_3, \cdots, x_n$ be a basis for $\mathfrak{M}$ with $d, x_2, \cdots, x_n$ a basis for $\mathfrak{L}$.

THEOREM 6. *The u-algebra of a nilpotent Lie algebra with a finite basis is a nilpotent associative algebra.*

The theorem is trivial if $\mathfrak{L}$ has 1 dimension. Suppose it true for algebras of order $n-1$. Choose $\mathfrak{M}$ and $d, x_2, \cdots, x_n$ as indicated. Then the u-algebra $\mathfrak{U}$ is generated by $d, x_2, \cdots, x_n$ and the u-algebra $\mathfrak{B}$ of $\mathfrak{M}$ is generated by $x_2, \cdots, x_n$. $\mathfrak{B}$ is nilpotent. The elements of $\mathfrak{U}$ have the form

$$u = v_0 + dv_1 + \cdots + d^{p-1}v_{p-1} + d\beta_1 + d^2\beta_2 + \cdots + d^{p-1}\beta_{p-1}$$

where $v_i \in \mathfrak{B}$ and $\beta_i \in \Phi$. The *weight* of $d^i v_1^{(\kappa_1)} v_2^{(\kappa_2)} \cdots v_s^{(\kappa_s)}$ where

$$v^{(\kappa)} = [\cdots [\overbrace{vd]d \cdots d}^{\kappa}] \in \mathfrak{B}$$

is defined to be $\geqq i+s+\kappa_1+\cdots+\kappa_s$. Hence the weight of each term of u is $\geqq 1$. Since

$$v_1^{(\kappa_1)} v_2^{(\kappa_2)} \cdots v_s^{(\kappa_s)} d = dv_1^{(\kappa_1)} v_2^{(\kappa_2)} \cdots v_s^{(\kappa_s)} + v_1^{(\kappa_1+1)} v_2^{(\kappa_2)} \cdots v_s^{(\kappa_s)} + \cdots + v_1^{(\kappa_1)} v_2^{(\kappa_2)} \cdots v_s^{(\kappa_s+1)}$$

the weight of a product $u_1u_2 \cdots u_k$ is $\geqq k$. If the index of nilpotency of $\mathfrak{B}$ is m and $d^{p^t}=0$, then every term of weight $\geqq mp^t$ is 0. Thus $\mathfrak{U}$ is nilpotent of index $\leqq mp^t$.

A consequence of this theorem is that the irreducible representations of a nilpotent restricted Lie algebra are all 0. Hence any representation has matrices in triangular form with diagonal elements 0 if the basis is properly chosen.

BIBLIOGRAPHY

1. I. Ado, Bulletin de la Société Physico-mathématique de Kazan, vol. 6 (1935).

2. G. Birkhoff, *Representability of Lie algebras* · · ·, Annals of Mathematics, (2), vol. 38 (1937), pp. 326–332.

3. N. Jacobson, *Abstract derivation and Lie algebras*, these Transactions, vol. 42 (1937), pp. 206–224.

4. O. Ore, *On a special class of polynomials*, these Transactions, vol. 35 (1933), pp. 559–584.

5. E. Witt, *Treue Darstellung Liescher Ringe*, Journal für die reine und angewandte Mathematik, vol. 177 (1937), pp. 152–160.

6. H. Zassenhaus, *Über Liesche Ringe mit Primzahlcharakteristik*, Abhandlungen aus dem mathematischen Seminar der Hansischen Universität, 1939, pp. 1–100.

7. ———, *Endliche p-Gruppe und Lie-Ring mit der Charakteristic p*, ibid., pp. 200–207.

8. M. Zorn, *On a theorem of Lie*, Bulletin of the American Mathematical Society, vol. 42 (1936), p. 485.

UNIVERSITY OF NORTH CAROLINA,
CHAPEL HILL, N. C.

CONFORMAL UNITARY SPACES

BY

N. COBURN

I. **Introduction.** In this paper, we shall study the relations between unitary spaces of n dimensions K_n and $'K_n$ whose fundamental tensors are related by the conformal transformation.

In §II, we develop the notation and some of the fundamental concepts of hermitian geometry. Section III deals with the relation between the paths of conformal unitary spaces. It is shown that (1) *two unitary spaces, both without torsion (symmetric connection), cannot be conformal*; (2) *if two unitary spaces are conformal, then their paths correspond*; (3) *conversely, if two unitary spaces, one of which is without torsion, are in a restricted correspondence of paths and if an unitary orthogonal ennuple of K_n corresponds conformally to an ennuple of $'K_n$ (see Definition 2), and if a certain curvature affinor $'C_{\alpha\mu\cdots}^{\cdots\beta}$ vanishes, then the two spaces are conformal.*

The paper concludes with a study of some particular conformal unitary spaces. We define the terms: (1) unitary k-space; (2) k-spaces which are k-conformal. We then prove the theorems: (1) *the affinor $'C_{\alpha\mu\cdots}^{\cdots\beta}$ is a conformal invariant of all k-spaces which are k-conformal*; (2) *if two unitary spaces are conformal and the affinor $'C_{\alpha\mu\cdots}^{\cdots\beta}$ is a conformal invariant, then the spaces are k-spaces.* A particular type of k-space with symmetric connection has been studied by Mitrochin, Fuchs, Bergman and others.

II. **Notation**[1]. Consider the "points" determined by assigning arbitrary values to the complex quantities ξ^λ (x^λ, y^λ real)

$$\xi^\lambda = x^\lambda + iy^\lambda, \qquad i = (-1)^{1/2}, \lambda = 1, \cdots, n. \tag{2.1}$$

These points build a complex space of n dimensions X_n (real topological X_{2n})[2]. Associated with (2.1) are the conjugate quantities

$$(\xi^\lambda)^* = \xi^{\lambda^*} = x^\lambda - iy^\lambda, \qquad \lambda^* = 1, \cdots, n. \tag{2.2}$$

Let us denote partial derivatives by

$$\partial_\mu = \partial/\partial\xi^\mu, \qquad \partial_{\mu^*} = \partial/\partial\xi^{\mu^*}; \tag{2.3}$$

then from (2.1) by the composite function theorem[3], we find for functions

Presented to the Society, February 24, 1940; received by the editors June 2, 1940.

(1) Since most of the references are to *Einführung in die neueren Methoden der Differentialgeometrie* by J. A. Schouten and D. J. Struik; vols. 1, 2; P. Noordhoff; Groningen, Batavia; 1935, 1938; we shall merely indicate volume, page and equation number.

(2) II, pp. 225–251.

(3) The equations (2.4) are obtained as follows:

of ξ^λ and ξ^{λ^*} respectively,

$$\partial_\mu\phi(\xi^\lambda) = \tfrac{1}{2}(\partial\phi/\partial x^\mu - i\partial\phi/\partial y^\mu),$$
$$\partial_{\mu^*}\psi(\xi^{\lambda^*}) = \tfrac{1}{2}(\partial\psi/\partial x^\mu + i\partial\psi/\partial y^\mu). \tag{2.4}$$

Consider the allowable coordinate transformations with nonvanishing Jacobian

$$\xi^\lambda = \xi^\lambda(\xi^{\lambda'}), \qquad \xi^{\lambda'} = x^{\lambda'} + iy^{\lambda'}. \tag{2.5}$$

The corresponding conjugate equations are

$$\xi^{\lambda^*} = \xi^{\lambda^*}(\xi^{\lambda^{*'}}), \qquad \xi^{\lambda^{*'}} = x^{\lambda'} - iy^{\lambda'}. \tag{2.6}$$

We now introduce the unit affinor whose intermediary(4) components are

$$A^\lambda_{\mu'} = \partial\xi^\lambda/\partial\xi^{\mu'}, \qquad A^{\lambda^*}_{\mu^{*'}} = \partial\xi^{\lambda^*}/\partial\xi^{\mu^{*'}}. \tag{2.7}$$

Evidently, if a star (*) after a symbol denotes the conjugate (i.e., replacing ξ^λ by ξ^{λ^*} and i by $-i$), then

$$(A^\lambda_{\mu'})^* = A^{\lambda^*}_{\mu^{*'}}. \tag{2.8}$$

This is a general notation which we will employ in the case of any affinor (also for the connection (2.15)).

We next consider a function of position in X_n, $\theta(\xi^\lambda, \xi^{\lambda^*})$, which is analytic(5) in ξ^λ, ξ^{λ^*} considered as independent variables. Such a function is called "semi-analytic." On the other hand, an analytic function $\phi(\xi^\lambda)$ of ξ^λ alone is called "analytic"; all the functions studied in this paper will be assumed to be "semi-analytic" or "analytic." The question arises as to the value of $\partial\xi^{\lambda^*}/\partial\xi^\lambda$. As is easily seen, such a derivative depends on the ratio dy^λ/dx^λ. In order to eliminate such a possibility, it will be desirable to consider ξ^{λ^*}, ξ^λ as independent variables, i.e.,

$$\partial_\mu\xi^{\lambda^*} = 0, \qquad \partial_{\mu^*}\xi^\lambda = 0. \tag{2.9}$$

We now formulate a definition of the operators ∂_μ, ∂_{μ^*} which will check with (2.9) and (2.4) and which will replace any previous conceptions of these operators.

$$\frac{\partial\phi}{\partial x^\mu} = \frac{\partial\phi}{\partial\xi^\alpha}\frac{\partial\xi^\alpha}{\partial x^\mu} = \partial_\mu\phi; \qquad \frac{\partial\phi}{\partial y^\mu} = \frac{\partial\phi}{\partial\xi^\alpha}\frac{\partial\xi^\alpha}{\partial y^\mu} = i\partial_\mu\phi.$$

By multiplying the second equation by $(-i)$ and adding these two equations, the first equation of (2.4) is obtained. See *Über unitäre Geometrie* by J. A. Schouten und D. van Dantzig, Mathematische Annalen, vol. 103 (1930), p. 324, footnote.

(4) I, p. 22.

(5) *On the elementary solution* · · · , T. Y. Thomas and E. W. Titt, Journal de Mathématique, vol. 18 (1939), p. 218 for definition of analytic "real" function. We assume a corresponding definition for analytic "complex" function.

FUNDAMENTAL DEFINITION. *For any "semi-analytic" function θ (and for the "analytic" functions ϕ, ψ), the operators $\partial_\mu, \partial_{\mu^*}$ will be defined by*

$$\partial_\mu \theta = \tfrac{1}{2}(\partial\theta/\partial x^\mu - i\partial\theta/\partial y^\mu), \tag{2.10a}$$

$$\partial_{\mu^*} \theta = \tfrac{1}{2}(\partial\theta/\partial x^\mu + i\partial\theta/\partial y^\mu). \tag{2.10b}$$

Replacing θ by ξ^{λ^*} in (2.10a) and using (2.2), we see that the first equality of (2.9) is valid. The second equality of (2.9) is obtained by replacing θ in (2.10b) by ξ^λ. Hence the definitions (2.10) are consistent with the equations (2.9). Thus ξ^λ, ξ^{λ^*} are independent variables with respect to differentiation. We note that $\partial(\xi^{\lambda^*})^n/\partial\xi^\lambda = 0$, identically in x^λ, y^λ.

One further question remains to be answered. "Is (2.10) identically satisfied in x^λ, y^λ?" The answer is in the affirmative. We give a proof for a semi-analytic function of one complex variable and its conjugate $\theta(z, z^*)$. Since θ is semi-analytic, we can expand θ in a convergent power series of terms $z^n z^{*m}$. Applying the operator ∂_* to this term, we obtain

$$\partial(z^n z^{*m})/\partial z = n z^{n-1} z^{*m}. \tag{2.10c}$$

Now if the expression $z^n z^{*m}$ is written as $f(x, y)g(x, y)$ and substituted into the right-hand side of (2.10a), we obtain

$$(g/2)(\partial f/\partial x - i\partial f/\partial y) + (f/2)(\partial g/\partial x - i\partial g/\partial y). \tag{2.10d}$$

From our observation at the conclusion of the preceding paragraph, we see that the second set of terms in (2.10d) vanishes. Furthermore, (2.10a) reduces to (2.4) (an identity) when $\phi = f = z^n$. That is,

$$\tfrac{1}{2}(\partial f/\partial x - i\partial f/\partial y) = n z^{n-1}. \tag{2.10e}$$

Hence (2.10c) and (2.10d) coincide and (2.10a) is an identity in x, y. The argument can be generalized to any number of variables.

We can now see why an analytic function $\phi(\xi^\lambda)$ of ξ^λ alone is called "analytic." From (2.9), we have

$$\partial_{\mu^*} \phi = 0. \tag{2.11}$$

But upon expanding ϕ into real and imaginary parts ($\phi = u + iv$) and using the right-hand side of (2.10b), we find from (2.11)

$$\partial u/\partial x^\mu - \partial v/\partial y^\mu = 0, \qquad \partial v/\partial x^\mu + \partial u/\partial y^\mu = 0. \tag{2.12}$$

These are the generalized Cauchy-Riemann equations. The conjugate equation to (2.11) is

$$\partial_\mu \phi^* = 0.$$

Evidently, any equation implies a conjugate equation. In the future, we shall indicate the validity of the conjugate by the abbreviation "conj." It is to

be noted, on the formal side, that in passing to the conjugate, indices are starred. The star of a starred quantity removes the original star.

We now introduce the vector $v^{\lambda}(\xi^{\lambda},\xi^{\lambda^*})$ of the *first type* with the transformation law

$$v^{\lambda'} = A_{\lambda}^{\lambda'}\, v^{\lambda}. \tag{2.13}$$

Associated with each such vector is its conjugate, or vector of the *second type* v^{λ^*}, with transformation law

$$v^{\lambda^{*'}} = A_{\lambda^*}^{\lambda^{*'}}\, v^{\lambda^*}. \tag{2.14}$$

The theory can be extended to affinors of any mixed valence.

Let us introduce a connection in X_n by means of the n^3 quantities $\Gamma_{\mu\alpha}^{\lambda}$ which are functions of position. Then, we define the covariant differential of a contravariant vector by

$$\delta v^{\lambda} = dv^{\lambda} + \Gamma_{\mu\alpha}^{\lambda} v^{\alpha} d\xi^{\mu}, \text{ conj.} \tag{2.15}$$

The $\Gamma_{\mu\alpha}^{\lambda}$ transform in a well known manner(6) under (2.5). We now write for a covariant vector

$$\delta w_{\lambda} = dw_{\lambda} - \Gamma_{\mu\lambda}^{\alpha} w_{\alpha} d\xi^{\mu}, \text{ conj.} \tag{2.16}$$

By expanding the ordinary differential of a vector, we obtain

$$dv^{\lambda} = d\xi^{\mu}\partial_{\mu} v^{\lambda} + d\xi^{\mu^*}\partial_{\mu^*} v^{\lambda}, \text{ conj.} \tag{2.17}$$

If we define the covariant derivative of w_{λ}, v^{λ} by means of

$$\nabla_{\mu} v^{\lambda} = \partial_{\mu} v^{\lambda} + \Gamma_{\mu\alpha}^{\lambda} v^{\alpha}, \qquad \nabla_{\mu} w_{\lambda} = \partial_{\mu} w_{\lambda} - \Gamma_{\mu\lambda}^{\alpha} w_{\alpha}, \tag{2.18}$$

$$\nabla_{\mu^*} v^{\lambda} = \partial_{\mu^*} v^{\lambda}, \qquad \nabla_{\mu^*} w_{\lambda} = \partial_{\mu^*} w_{\lambda}, \text{ conj.}, \tag{2.19}$$

then

$$\delta v^{\lambda} = d\xi^{\mu}\nabla_{\mu} v^{\lambda} + d\xi^{\mu^*}\nabla_{\mu^*} v^{\lambda}, \text{ conj.}, \tag{2.20}$$

$$\delta w_{\lambda} = d\xi^{\mu}\nabla_{\mu} w_{\lambda} + d\xi^{\mu^*}\nabla_{\mu^*} w_{\lambda}, \text{ conj.} \tag{2.21}$$

An hermitian X_n with a covariant derivative defined by (2.18), (2.19) is denoted by K_n.

We now introduce an hermitian tensor with hermitian symmetry

$$a_{\lambda\mu^*} = (a_{\lambda\mu^*})^{*\prime} = a_{\mu^*\lambda}, \tag{2.22}$$

the $(')$ indicating the transpose matrix. If we condition $a_{\lambda\mu^*}$ by requiring that

(6) I, p. 75, (7.4).

$$(2.23)\qquad \delta a_{\lambda\mu^*} = 0 = (\partial_\nu a_{\lambda\mu^*} - \Gamma^{\rho}_{\nu\lambda} a_{\rho\mu^*})d\xi^{\nu} + (\partial_{\nu^*} a_{\lambda\mu^*} - \Gamma^{\rho^*}_{\nu^*\mu^*} a_{\lambda\rho^*})d\xi^{\nu^*},$$

then the space K_n becomes a unitary K_n. For such a space, from (2.23), we can prove(7)

$$(2.24)\qquad \nabla_\nu a_{\lambda\mu^*} = \partial_\nu a_{\lambda\mu^*} - \Gamma^{\rho}_{\nu\lambda} a_{\rho\mu^*} = 0, \text{ conj.},$$

$$(2.25)\qquad \nabla_{\nu^*} a_{\lambda\mu^*} = \partial_{\nu^*} a_{\lambda\mu^*} - \Gamma^{\rho^*}_{\nu^*\mu^*} a_{\lambda\rho^*} = 0, \text{ conj.}$$

The $a_{\lambda\mu^*}$ is now a fundamental tensor and can be used to raise and lower indices through the ∇ operator; that is, if

$$(2.26)\qquad v_\lambda = a_{\lambda\mu^*} v^{\mu^*}, \text{ conj.},$$

then

$$(2.27)\qquad \nabla_\alpha v_\lambda = a_{\lambda\mu^*} \nabla_\alpha v^{\mu^*}, \text{ conj.}$$

If we define the contravariant fundamental tensor $a^{\sigma^*\kappa}$ by

$$(2.28)\qquad a^{\sigma^*\kappa} a_{\lambda\sigma^*} = A^{\kappa}_{\lambda}, \text{ conj.},$$

then (2.24), (2.25) may be solved for the connection

$$(2.29)\qquad \Gamma^{\kappa}_{\mu\lambda} = (\partial_\mu a_{\lambda\sigma^*}) a^{\sigma^*\kappa} = (\partial_\mu a_{\sigma^*\lambda}) a^{\kappa\sigma^*},$$

$$(2.30)\qquad \Gamma^{\kappa^*}_{\mu^*\lambda^*} = (\partial_{\mu^*} a_{\sigma\lambda^*}) a^{\kappa^*\sigma} = (\partial_{\mu^*} a_{\lambda^*\sigma}) a^{\sigma\kappa^*}.$$

For the curvature affinor, we have the components(8),

$$(2.31)\qquad R^{\cdots\kappa}_{\nu\mu\lambda} = 2\partial_{[\nu}\Gamma^{\kappa}_{\mu]\lambda} + 2\Gamma^{\kappa}_{[\nu|\rho|}\Gamma^{\rho}_{\mu]\lambda}, \text{ conj.},$$

$$(2.32)\qquad R^{\cdots\kappa}_{\nu^*\mu\lambda} = \partial_{\nu^*}\Gamma^{\kappa}_{\mu\lambda}, \text{ conj.}$$

The sign [ ] on the indices means that the alternating product is to be formed; the sign ( ) enclosing indices means that the symmetric product is to be formed; the sign | | implies that indices enclosed are not to be included in any alternating or symmetric product(9).

For unitary K_n, we find from (2.29)(10)

$$(2.33)\qquad R^{\cdots\kappa}_{\nu\mu\lambda} = 0, \text{ conj.}$$

Next, we introduce the torsion affinor

$$(2.34)\qquad S^{\cdots\kappa}_{\mu\lambda} = \Gamma^{\kappa}_{[\mu\lambda]}, \text{ conj.}$$

---

(7) II, p. 234, (21.40).
(8) II, p. 233, (21.33).
(9) I, pp. 14, 15.
(10) II, p. 235, (21.43).

If this affinor does not vanish in K_n, then we may say that K_n has torsion.

Finally, we introduce an unitary-orthogonal ennuple[11] at each point P of K_n. We may denote these vectors by $u^\lambda(\xi^\lambda, \xi^{\lambda^*})$ (subindex $k=1, \cdots, n$) where

$$a_{\lambda a^*} u_k^\lambda u_j^{a^*} = \delta_{kj}; \qquad u_k^\lambda u_\lambda^j = \delta_k^j, \text{ conj.} \tag{2.35}$$

Here δ is the Kronecker symbol. We note that at each point P there are $\infty^{n(n-1)/2}$ such ennuples. Let $'K_n$ be another unitary space also mapped by the variables ξ^λ. We say that the point $P(\xi^\lambda)$ of K_n corresponds to the point $'P(\xi^\lambda)$ of $'K_n$ and conversely. Let $'u^\lambda(\xi^\lambda, \xi^{\lambda^*})$, subindex $k=1, \cdots, n$, denote an unitary orthogonal ennuple at $'P(\xi^\lambda)$. We now form the definitions

DEFINITION 1. *We shall say that two ennuples of K_n and $'K_n$ at corresponding points P and $'P$ are corresponding ennuples if functions $\alpha(\xi^\lambda, \xi^{\lambda^*})$, subindex $j=1, \cdots, n$, exist so that*

$$'u_j^\lambda = \alpha_j\, 'u_j^\lambda, \qquad \textit{not summed on } j. \tag{2.36}$$

DEFINITION 2. *If the α (subindex j) are all equal, we say that the two ennuples correspond conformally.*

III. **Conformal unitary spaces; correspondence of paths.** Consider two unitary spaces K_n, $'K_n$ both mapped by the variables ξ^λ and whose fundamental tensors are related by

$$'a_{\lambda\mu^*} = \sigma a_{\lambda\mu^*}, \text{ conj.} \tag{3.1}$$

Here σ is a function of ξ^λ, ξ^{λ^*}. Taking the conjugate of (3.1) and using (2.22), we see that σ is real. Hence,

DEFINITION 3. *Two unitary spaces whose fundamental tensors are related by* (3.1) *where σ is real will be said to be conformal.*

By use of (3.1) and (2.29), we find

$$'\Gamma_{\mu\lambda}^\kappa = \Gamma_{\mu\lambda}^\kappa + A_\lambda^\kappa \partial_\mu \log \sigma, \text{ conj.} \tag{3.2}$$

We shall now prove the theorem

THEOREM 1. *Unitary spaces K_n, $'K_n$, both without torsion, cannot be conformal, unless K_n and $'K_n$ are the same space.*

For if these spaces were conformal, then (3.2) would be valid. By our assumption, the connections of both spaces are symmetric. Thus

---

[11] I, p. 235.

$$A^{\kappa}_{[\mu\partial_\lambda]} \log \sigma = 0, \text{ conj.} \tag{3.3}$$

By expanding, we find

$$\partial_\mu \log \sigma = 0, \text{ conj.} \tag{3.4}$$

Hence σ is a constant. Such a conformal transformation takes K_n into itself.

We shall assume in the future that at least one of the unitary spaces K_n, $'K_n$ has torsion. If only one of these spaces has torsion (K_n), then from (2.35), (3.2) we see that

$$S^{\cdot\cdot\kappa}_{\mu\lambda} = -A^{\kappa}_{[\mu p_\lambda]}, \qquad p_\lambda = \partial_\lambda \log \sigma, \text{ conj.} \tag{3.5}$$

The connection of K_n is then said to be semi-symmetric[12]. In the general case, we may write (3.2) in the form

$$'\Gamma^{\kappa}_{(\mu\lambda)} = \Gamma^{\kappa}_{(\mu\lambda)} + A^{\kappa}_{(\mu p_\lambda)}, \text{ conj.} \tag{3.6}$$

If we denote these symmetric parts of the connection by

$$'\mathring{\Gamma}^{\kappa}_{\mu\lambda} = {}'\Gamma^{\kappa}_{(\mu\lambda)}, \qquad \mathring{\Gamma}^{\kappa}_{\mu\lambda} = \Gamma^{\kappa}_{(\mu\lambda)}, \text{ conj.} \tag{3.7}$$

and replace the vector p_λ by

$$q_\lambda = p_\lambda/2 = \partial_\lambda \log \sigma/2, \text{ conj.}; \tag{3.8}$$

then (3.6) becomes

$$'\mathring{\Gamma}^{\kappa}_{\mu\lambda} = \mathring{\Gamma}^{\kappa}_{\mu\lambda} + 2A^{\kappa}_{(\mu q_\lambda)}, \text{ conj.} \tag{3.9}$$

But, the equations (3.9) are well known in the theory of paths. These equations state that the paths of K_n and $'K_n$ correspond[13]. Hence, we have the theorem:

THEOREM 2. *If two unitary spaces are conformal, then their paths correspond.*

This last theorem has no equivalent in the theory of real spaces.

We shall now prove a converse of Theorem 2. Consider two unitary spaces: $'K_n$ without torsion; K_n with torsion. Let us assume that the paths of these spaces correspond. Hence the connections of these spaces are related by

$$'\Gamma^{\kappa}_{\mu\lambda} = \mathring{\Gamma}^{\kappa}_{\mu\lambda} + 2A^{\kappa}_{(\mu q_\lambda)}, \text{ conj.}, \tag{3.10}$$

where

$$\mathring{\Gamma}^{\kappa}_{\mu\lambda} = \Gamma^{\kappa}_{(\mu\lambda)}, \text{ conj.} \tag{3.11}$$

---

[12] I, p. 82, (7.27).
[13] II, p. 176, (18.1).

In our work, we shall find it useful to consider the particular correspondence of paths for which $q_\lambda(\xi^\lambda, \xi^{\lambda^*})$ is a gradient vector. Hence, we define:

DEFINITION 4. *The particular correspondence of paths determined by* (3.10) *when q_λ is a gradient vector of a real scalar function of position shall be called restricted correspondence of paths.*

The curvature affinors determined by the connections occurring in (3.10) may be found from (2.31), (2.32). With the aid of (3.10), we find that these affinors are related by[14]

$$'R_{\nu\mu\lambda}^{\cdots\kappa} = \overset{s}{R}_{\nu\mu\lambda}^{\cdots\kappa} + 2A^{\kappa}_{[\nu}q_{\mu]\lambda} - 2q_{[\nu\mu]}A^{\kappa}_{\lambda}, \text{ conj.}, \tag{3.12}$$

$$'R_{\nu^*\mu\lambda}^{\cdots\kappa} = \overset{s}{R}_{\nu^*\mu\lambda}^{\cdots\kappa} + 2(\overset{s}{\nabla}_{\nu^*}q_{(\mu)}A^{\kappa}_{\lambda)}, \text{ conj.}, \tag{3.13}$$

where

$$q_{\nu\mu} = -\overset{s}{\nabla}_\nu q_\mu + q_\nu q_\mu, \text{ conj.} \tag{3.14}$$

In these equations, the ∇_μ (superindex s) represents covariant differentiation with respect to the connection occurring in the right-hand side of (3.10). Furthermore, by (2.33), the left-hand side of (3.12) vanishes. Finally, we define the affinors

$$'R_{\alpha\lambda^*\cdot\cdot\nu}^{\cdot\cdot\beta} = 'R_{\alpha\mu^*\lambda^*}^{\cdots\kappa^*}\,'a^{\lambda^*\beta}\,'a_{\nu\kappa^*}, \text{ conj.}, \tag{3.15}$$

$$'C_{\alpha\mu^*\nu}^{\cdots\beta} = 'R_{\alpha\mu^*\cdot\cdot\nu}^{\cdot\cdot\beta} - 'R_{\mu^*\alpha\nu}^{\cdots\beta}, \text{ conj.} \tag{3.16}$$

We may now state our theorem as:

THEOREM 3. *If two unitary spaces, $'K_n$ without torsion, K_n with torsion are given, and if* (1) *the paths of these spaces are in restricted correspondence,* (2) *the affinor $'C_{\alpha\mu^*\nu}^{\cdots\beta}$ vanishes,* (3) *one arbitrary unitary orthogonal ennuple at one arbitrary point P of K_n is conformal to the corresponding ennuple at the corresponding point $'P$ of $'K_n$* (*see* (2.36)), *then the two spaces are conformal.*

By assumption (1), the equation (3.10) is valid. Substituting (2.29) into (3.10) and transvecting with $'a_{\kappa\alpha^*}$, we find

$$\partial_\mu\,'a_{\lambda\alpha^*} = \overset{s}{\Gamma}{}^{\kappa}_{\mu\lambda}\,'a_{\kappa\alpha^*} + 'a_{\mu\alpha^*}q_\lambda + 'a_{\lambda\alpha^*}q_\mu, \text{ conj.} \tag{3.17}$$

We shall consider (3.17) as a system of partial differential equations in $'a_{\lambda\alpha^*}$ where $\Gamma^{\kappa}_{\mu\lambda}$ (superindex s) and q_λ are considered as known. The integrability conditions of (3.17) are

$$\partial_{[\nu}\partial_{\mu]}\,'a_{\lambda\alpha^*} = 0, \text{ conj.}, \tag{3.18}$$

$$\partial_{[\nu^*}\partial_{\mu]}\,'a_{\lambda\alpha^*} = 0, \text{ conj.} \tag{3.19}$$

---

[14] II, p. 240, (21.62).

Our method of procedure is to show that because of our assumptions, the equations (3.18), (3.19) are identically satisfied. It then follows that there exists an unique solution for $'a_{\lambda\mu^*}$ which by (2.22) contains $n(n+1)/2$ constants of integration. From (3.11), (2.29) the equation (3.17) may be written in the form

$$\partial_{(\mu}{}'a_{\lambda)\alpha^*} = (\partial_{(\mu}a_{\lambda)\beta^*})a^{\beta^*\varepsilon}\,{}'a_{\varepsilon\alpha^*} + {}'a_{\mu\alpha^*}q_\lambda + {}'a_{\lambda\alpha^*}q_\mu, \text{ conj.} \tag{3.20a}$$

Since the connection of $'K_n$ is symmetric, (3.20) may also be written in the form

$$\partial_\mu{}'a_{\lambda\alpha^*} = (\partial_{(\mu}a_{\lambda)\beta^*})a^{\beta^*\varepsilon}\,{}'a_{\varepsilon\alpha^*} + {}'a_{\mu\alpha^*}q_\lambda + {}'a_{\lambda\alpha^*}q_\mu, \text{ conj.} \tag{3.20b}$$

This last set of equations constitutes a system of linear homogeneous partial differential equations. For, we can write (3.20b) in the form

$$\partial_\mu\,{}'a_{\lambda\alpha^*} - A_{\mu\lambda}^{\cdots\nu}\,{}'a_{\nu\alpha^*} = 0, \text{ conj.}, \tag{3.20c}$$

where $A_{\mu\lambda}^{\cdots\nu}$ are known functions of ξ^λ defined by

$$A_{\mu\lambda}^{\cdots\nu} = a^{\beta^*\nu}(\partial_{(\mu}a_{\lambda)\beta^*}) + A_\mu^{\nu}q_\lambda + A_\lambda^{\nu}q_\mu, \text{ conj.} \tag{3.20d}$$

If $n(n+1)/2$ particular integrals of (3.20c) are denoted by

$$'^{k}_{a_{\lambda\alpha^*}}, \qquad k = 1, \cdots, n(n+1)/2, \tag{3.21a}$$

then the general solution of (3.20c) is given by

$$'a_{\lambda\alpha^*} = c_k\,{}'^{k}_{a_{\lambda\alpha^*}}, \quad k = 1, \cdots, n(n+1)/2; \text{ conj.}; \tag{3.21b}$$

where c, subindex k, are arbitrary complex constants. By substitution, we see that one particular solution of (3.20a) or (3.20c) is

$$'^{1}_{a_{\lambda\alpha^*}} = \sigma a_{\lambda\alpha^*}, \text{ conj.}, \tag{3.21c}$$

where σ is a real solution of

$$\partial_\mu \log \sigma = 2q_\mu, \text{ conj.} \tag{3.21d}$$

Since q_μ is the gradient of a real scalar ϕ, the solution of (3.21d) for σ is given by

$$\log \sigma = 2\phi + f(\xi^{\lambda^*}) + s, \tag{3.21e}$$

where s is an arbitrary constant and f is an arbitrary function. The conjugate equation of (3.21d) yields

$$\log \sigma = 2\phi + g(\xi^\lambda) + r. \tag{3.21f}$$

This last result follows from the facts that σ, ϕ are real; r is an arbitrary con-

stant and g is an arbitrary function. Upon taking the conjugate of (3.21e), we find

$$s^* + f^*(\xi^\lambda) = g^*(\xi^{\lambda^*}) + r^*. \tag{3.21g}$$

Since ξ^λ, ξ^{λ^*} are independent, the functions f, g both vanish. Furthermore, $s=r$ and s is real. Thus σ is given by

$$\sigma = Se^{2i\phi}; \tag{3.21h}$$

real S is a real arbitrary constant.

The general solution of (3.20c) can now be written as

$$'a_{\lambda\alpha^*} = \sigma a_{\lambda\alpha^*} + \underset{k}{c}\, {}'\overset{k}{a}_{\lambda\alpha^*}, \quad k = 2, \cdots, n(n+1)/2; \text{ conj.} \tag{3.22a}$$

Note, that σ contains an arbitrary constant. We now use condition (3) of our theorem to show that c (subindex $k=2, \cdots$) must be zero. At P, we expand $'a_{\lambda\alpha^*}$ (superindex k) in terms of the ennuple and find

$${}'\overset{k}{a}_{\lambda\alpha^*} = \underset{jl}{\overset{k}{a}}\, \overset{j}{u}_\lambda \overset{l}{u}_{\alpha^*}, \text{ conj.} \tag{3.22b}$$

By substituting (3.22b) into (3.22a) and transvecting with the u_λ, u_{λ^*} vectors, we obtain by use of (2.35) and (2.36)—where all α's are equal—

$$\underset{k}{c}\, \underset{jl}{\overset{k}{a}} = 0, \qquad j \neq l,\ j, l = 1, \cdots, n, \tag{3.22c}$$

$$\underset{k}{c}\, \underset{jj}{\overset{k}{a}} = (\alpha\alpha^* - \sigma), \quad k = 2, \cdots, n(n+1)/2. \tag{3.22d}$$

We can evidently fix the arbitrary constant in σ so that the right-hand side of (3.22d) vanishes at some point P. Let us assume that the c (subindex k) are not all zero. Then the $[n(n+1)/2-1]$ solutions $'a_{\lambda\mu^*}$ (superindex k) are linearly related at P. Since the $'a_{\lambda\mu^*}$ (superindex k) satisfy a system of linear partial differential equations, these $'a_{\lambda\mu^*}$ are not independent in K_n. It will be shown in (3.41) that the integrability conditions of (3.20a) are identically satisfied in virtue of condition (2) of our theorem. Hence, from the theory of linear differential equations, (3.20a) has $n(n+1)/2$ independent integrals $\sigma a_{\lambda\mu^*}$, $'a_{\lambda\mu^*}$ (superindex k). Thus c (subindex k) must vanish and the only solution of (3.20a) under the condition (3) of our theorem is

$$'a_{\lambda\alpha^*} = \sigma a_{\lambda\alpha^*}, \text{ conj.} \tag{3.23}$$

Thus, we have reduced our problem to showing that the conditions (3.18), (3.19) are identically satisfied. We first consider (3.18). From (3.17), we find

$$\begin{aligned} \partial_{[\nu}\partial_{\mu]}\, {}'a_{\lambda\alpha^*} = {}& {}'a_{\kappa\alpha^*}(\partial_{[\nu}\overset{\kappa}{\Gamma}_{\mu]\lambda}) + \overset{\kappa}{\Gamma}_{\lambda[\mu}\partial_{\nu]}\, {}'a_{\kappa\alpha^*} + q_\lambda \partial_{[\nu}\, {}'a_{\mu]\alpha^*} \\ & + q_{[\mu}\partial_{\nu]}\, {}'a_{\lambda\alpha^*} + (\partial_{[\nu}q_{|\lambda|})'a_{\mu]\alpha^*} + {}'a_{\lambda\alpha^*}\partial_{[\nu}q_{\mu]}, \text{ conj.} \end{aligned} \tag{3.24}$$

With the aid of (3.17), we see that the third term on the right-hand side of (3.24) vanishes. We next simplify terms two and four:

$$\overset{\circ}{\Gamma}{}^{\kappa}_{\lambda[\mu}\partial_{\nu]}{}'a_{\kappa\alpha^*} = \overset{\circ}{\Gamma}{}^{\kappa}_{\lambda[\mu}\overset{\circ}{\Gamma}{}^{\rho}_{\nu]\kappa}{}'a_{\rho\alpha^*} + \overset{\circ}{\Gamma}{}^{\kappa}_{\lambda[\mu}{}'a_{\nu]\alpha^*}q_{\kappa} + \overset{\circ}{\Gamma}{}^{\kappa}_{\lambda[\mu}q_{\nu]}{}'a_{\kappa\alpha^*}, \tag{3.25}$$

$$q_{[\mu}\partial_{\nu]}{}'a_{\lambda\alpha^*} = \overset{\circ}{\Gamma}{}^{\kappa}_{\lambda[\nu}q_{\mu]}{}'a_{\kappa\alpha^*} + q_{[\mu}{}'a_{\nu]\alpha^*}q_{\lambda} + q_{[\mu}q_{\nu]}{}'a_{\lambda\alpha^*}, \text{ conj.} \tag{3.26}$$

Upon substituting these simplifications into (3.24) and using (2.31), (3.14), we obtain

$$\partial_{[\nu}\partial_{\mu]}{}'a_{\lambda\alpha^*} = \tfrac{1}{2}\overset{\circ}{R}{}^{\cdots\kappa}_{\nu\mu\lambda}{}'a_{\kappa\alpha^*} + q_{\lambda[\mu}{}'a_{\nu]\alpha^*} - q_{[\nu\mu]}{}'a_{\lambda\alpha^*}, \text{ conj.} \tag{3.27}$$

From (3.12), (3.27) we find

$$2\,'a^{\alpha^*\beta}\partial_{[\nu}\partial_{\mu]}{}'a_{\lambda\alpha^*} = {}'R^{\cdots\beta}_{\nu\mu\lambda}, \text{ conj.} \tag{3.28}$$

Since the right-hand side of (3.28) vanishes for a unitary $'K_n$, we see that the conditions (3.18) are identically satisfied.

The verification of (3.19) requires considerable computation due to the unsymmetric character of the indices. We shall list the principal steps. By taking the conjugate of (3.9) and using (2.30), we find

$$\partial_{\mu^*}{}'a_{\nu\lambda^*} = \overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}{}'a_{\nu\kappa^*} + {}'a_{\nu\mu^*}q_{\lambda^*} + {}'a_{\nu\lambda^*}q_{\mu^*}, \text{ conj.} \tag{3.29}$$

By differentiation of (3.29),

$$\begin{aligned}\partial_{\alpha}\partial_{\mu^*}{}'a_{\nu\lambda^*} = {} & (\partial_{\alpha}\overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*})'a_{\nu\kappa^*} + (\partial_{\alpha}{}'a_{\nu\kappa^*})\overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*} + (\partial_{\alpha}{}'a_{\nu\mu^*})q_{\lambda^*} \\ & + {}'a_{\nu\mu^*}\partial_{\alpha}q_{\lambda^*} + (\partial_{\alpha}{}'a_{\nu\lambda^*})q_{\mu^*} + {}'a_{\nu\lambda^*}\partial_{\alpha}q_{\mu^*}, \text{ conj.}\end{aligned} \tag{3.30}$$

We simplify terms two, three and five of the right-hand side by means of (3.17),

$$\overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}\partial_{\alpha}{}'a_{\nu\kappa^*} = \overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}\overset{\circ}{\Gamma}{}^{\beta}_{\alpha\nu}{}'a_{\beta\kappa^*} + \overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}{}'a_{\alpha\kappa^*}q_{\alpha} + \overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}{}'a_{\alpha\kappa^*}q_{\nu}, \text{ conj.}, \tag{3.31}$$

$$(\partial_{\alpha}{}'a_{\nu\mu^*})q_{\lambda^*} = \overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}{}'a_{\kappa\mu^*}q_{\lambda^*} + {}'a_{\alpha\mu^*}q_{\nu}q_{\lambda^*} + {}'a_{\nu\mu^*}q_{\alpha}q_{\lambda^*}, \text{ conj.}, \tag{3.32}$$

$$(\partial_{\alpha}{}'a_{\nu\lambda^*})q_{\mu^*} = \overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}{}'a_{\kappa\lambda^*}q_{\mu^*} + {}'a_{\alpha\lambda^*}q_{\nu}q_{\mu^*} + {}'a_{\nu\lambda^*}q_{\alpha}q_{\mu^*}, \text{ conj.} \tag{3.33}$$

From (3.17), by differentiation, we get

$$\begin{aligned}\partial_{\mu^*}\partial_{\alpha}{}'a_{\nu\lambda^*} = {} & (\partial_{\mu^*}\overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu})'a_{\kappa\lambda^*} + \overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}\partial_{\mu^*}{}'a_{\kappa\lambda^*} + (\partial_{\mu^*}{}'a_{\alpha\lambda^*})q_{\nu} \\ & + {}'a_{\alpha\lambda^*}\partial_{\mu^*}q_{\nu} + (\partial_{\mu^*}{}'a_{\nu\lambda^*})q_{\alpha} + {}'a_{\nu\lambda^*}\partial_{\mu^*}q_{\alpha}, \text{ conj.}\end{aligned} \tag{3.34}$$

For the terms two, three and five of (3.34), we obtain

$$\overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}\partial_{\mu^*}{}'a_{\nu\lambda^*} = \overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}\overset{\circ}{\Gamma}{}^{\rho^*}_{\mu^*\lambda^*}{}'a_{\kappa\rho^*} + \overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}{}'a_{\nu\mu^*}q_{\lambda^*} + \overset{\circ}{\Gamma}{}^{\kappa}_{\alpha\nu}{}'a_{\kappa\lambda^*}q_{\mu^*}, \text{ conj.}, \tag{3.35}$$

$$(\partial_{\mu^*}{}'a_{\alpha\lambda^*})q_{\nu} = \overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}{}'a_{\alpha\kappa^*}q_{\nu} + {}'a_{\alpha\mu^*}q_{\lambda^*}q_{\nu} + {}'a_{\alpha\lambda^*}q_{\mu^*}q_{\nu}, \text{ conj.}, \tag{3.36}$$

$$(\partial_{\mu^*}{}'a_{\nu\lambda^*})q_{\alpha} = \overset{\circ}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*}{}'a_{\nu\lambda^*}q_{\alpha} + {}'a_{\nu\lambda^*}q_{\mu^*}q_{\alpha} + {}'a_{\nu\mu^*}q_{\lambda^*}q_{\alpha}, \text{ conj.} \tag{3.37}$$

When we subtract (3.30) from (3.34), we find that the set (3.31), (3.32), (3.33) cancels against (3.35), (3.36), (3.37). And we obtain

$$
\begin{aligned}
(3.38)\qquad 2\partial_{[\alpha}\partial_{\mu^*]}{}'a_{\nu\lambda^*} &= (\partial_\alpha \overset{*}{\Gamma}{}^{\kappa^*}_{\mu^*\lambda^*})'a_{\nu\kappa^*} - (\partial_{\mu^*}\overset{*}{\Gamma}{}^{\kappa}_{\alpha\nu})'a_{\kappa\lambda^*} \\
&\quad + {}'a_{\nu\mu^*}\partial_\alpha q_{\lambda^*} + {}'a_{\nu\lambda^*}\partial_\alpha q_{\mu^*} - {}'a_{\alpha\lambda^*}\partial_{\mu^*}q_\nu - {}'a_{\nu\lambda^*}\partial_{\mu^*}q_\alpha, \text{ conj.}
\end{aligned}
$$

This equation may be simplified by transvection and use of (2.32) into

$$
\begin{aligned}
(3.39)\qquad 2\,{}'a^{\lambda^*\beta}\partial_{[\alpha}\partial_{\mu^*]}{}'a_{\nu\lambda^*} &= \overset{*}{R}{}^{\cdots\kappa^*}_{\alpha\mu^*\lambda^*}{}'a^{\lambda^*\beta}{}'a_{\nu\kappa^*} - \overset{*}{R}{}^{\cdots\beta}_{\mu^*\alpha\nu} \\
&\quad + {}'a^{\lambda^*\beta}{}'a_{\nu\mu^*}\partial_\alpha q_{\lambda^*} + A^{\beta}_{\nu}\partial_\alpha q_{\mu^*} - A^{\beta}_{\alpha}\partial_{\mu^*}q_\nu - A^{\beta}_{\nu}\partial_{\mu^*}q_\alpha, \text{ conj.}
\end{aligned}
$$

Upon replacing the first and second terms of the right-hand side by their values as determined by (3.13), and using (3.15), we find

$$
(3.40)\qquad 2\,{}'a^{\lambda^*\beta}\partial_{[\alpha}\partial_{\mu^*]}{}'a_{\nu\lambda^*} = {}'R^{\cdots\beta}_{\alpha\mu^*\cdot\nu} - {}'R^{\cdots\beta}_{\mu^*\alpha\nu}, \text{ conj.}
$$

From (3.40), (3.16), these last integrability conditions are equivalent to

$$
(3.41)\qquad {}'C^{\cdots\beta}_{\alpha\mu^*\nu} = 0, \text{ conj.}
$$

Hence our theorem is proved.

The extension of this theorem to K_n, $'K_n$ both with torsion is not possible. For then the term $'S^{\cdots\kappa}_{\mu\lambda}{}'a_{\kappa\alpha^*}$ must be added to the right-hand side of (3.17); the argument (3.20, 3.21) is then no longer valid.

IV. **Conformal unitary k-spaces; a conformal invariant.** We first define a unitary k-space by

DEFINITION 5. *If the fundamental tensor of a unitary space K_n can be written in the form $a_{\alpha\nu^*}=\beta\partial_\alpha\partial_{\nu^*}\log k$, where k is real and β is an arbitrary function of ξ^λ, ξ^{λ^*} then K_n shall be said to be a unitary k-space.*

From (2.29), we see that those k-spaces for which $\beta=1$ have a symmetric connection. Such spaces have been studied by Mitrochin, Fuchs, Bergman and others([15]). The class of k-spaces with real β are evidently conformal to those k-spaces for which $\beta=1$ (fixed k). Hence by Theorem 1, the connections of the class of k-spaces with real β (β not constant) possess torsion.

If we fix the function k, but allow β to vary, then we obtain a set of such k-spaces. In particular, we consider those β which are the product of a real function and an initial β_1. Any two spaces of such a set are conformal to each other. We define a subset of such spaces by the following:

([15]) *Über die Veränderung der Krümmung von Hyperflächen bei Pseudokonformen Abbildungen*, by I. Mitrochin, Bulletin de l' Institute de Mathématiques et Mécanique à l' Université Koubycheff de Tomsk, vol. 1 (1937), no. 3.

DEFINITION 6. *All unitary k-spaces (fixed k) whose fundamental tensors are related by*

$$'a_{\lambda\mu^*} = k^\alpha a_{\lambda\mu^*}, \text{ conj.}, \tag{4.1}$$

where α is any real number, will be said to be k-conformal.

This definition imposes the following restrictions on β. Consider two k-spaces

$$a_{\lambda\mu^*} = \beta_1 \partial_\lambda \partial_{\mu^*} \log k, \text{ conj.}, \tag{4.2}$$

$$'a_{\lambda\mu^*} = \beta_2 \partial_\lambda \partial_{\mu^*} \log k, \text{ conj.}. \tag{4.3}$$

If these k-spaces are k-conformal, then

$$k^\alpha \beta_1 = \beta_2. \tag{4.4}$$

Evidently, the ∞^1 k-spaces whose first fundamental tensors are

$$\beta_1 \partial_\lambda \partial_{\mu^*} \log k; \quad k^\alpha \beta_1 \partial_\lambda \partial_{\mu^*} \log k; \quad k^{\alpha'} \beta_1 \partial_\lambda \partial_{\mu^*} \log k, \text{ etc.}, \tag{4.5}$$

are k-conformal to each other when α, α', etc., are any real numbers. We now prove the theorem

THEOREM 4. *The affinor $C^{\cdots\beta}_{\alpha\mu^*\nu}$ is a conformal invariant of all unitary k-spaces which are k-conformal.*

From (3.16), (3.10), (2.32), we find for any two unitary conformal spaces

$$'C^{\cdots\beta}_{\alpha\mu^*\nu} = {}'R^{\cdots\varepsilon^*}_{\alpha\mu^*\lambda^*} {}'a^{\lambda^*\beta}\, {}'a_{\nu\varepsilon^*} - {}'R^{\cdots\beta}_{\mu^*\alpha\nu}, \text{ conj.}, \tag{4.6}$$

$$\begin{aligned} 'C^{\cdots\beta}_{\alpha\mu^*\nu} &= (\partial_\alpha \overset{\varepsilon}{\Gamma}{}^{\varepsilon^*}_{\mu^*\lambda^*} + A^{\varepsilon^*}_{\mu^*} \partial_\alpha q_{\lambda^*} + A^{\varepsilon^*}_{\lambda^*} \partial_\alpha q_{\mu^*}) {}'a^{\lambda^*\beta}\, {}'a_{\nu\varepsilon^*} \\ &\quad - (\partial_{\mu^*} \overset{\varepsilon}{\Gamma}{}^{\beta}_{\alpha\nu} + A^{\beta}_{\alpha} \partial_{\mu^*} q_\nu + A^{\beta}_{\nu} \partial_{\mu^*} q_\alpha), \text{ conj.} \end{aligned} \tag{4.7}$$

Simplifying (4.7) with the aid of (3.16), we obtain

$$\begin{aligned} 'C^{\cdots\beta}_{\alpha\mu^*\nu} &= C^{\cdots\beta}_{\alpha\mu^*\nu} + {}'a^{\lambda^*\beta}\, {}'a_{\nu\mu^*} \partial_\alpha q_{\lambda^*} + A^{\beta}_{\nu} \partial_\alpha q_{\mu^*} \\ &\quad - A^{\beta}_{\alpha} \partial_{\mu^*} q_\nu - A^{\beta}_{\nu} \partial_{\mu^*} q_\alpha, \text{ conj.} \end{aligned} \tag{4.8}$$

Since our spaces are k-spaces and k-conformal, then (4.2), (4.3), (4.4) are valid. Furthermore, from (3.8)

$$q_{\lambda^*} = \tfrac{1}{2} \partial_{\lambda^*} \log \sigma, \text{ conj.} \tag{4.9}$$

Hence from (4.1), where $\sigma = k^\alpha$, and (4.3), we find

$$\partial_\mu q_{\lambda^*} = (\alpha/2) \partial_\mu \partial_{\lambda^*} \log k = \alpha\, 'a_{\mu\lambda^*}/2\beta_2, \text{ conj.} \tag{4.10}$$

From (4.9) conj., we have

$$\partial_{\lambda^*} q_\mu = (\alpha/2) \partial_{\lambda^*} \partial_\mu \log k = \alpha\, 'a_{\mu\lambda^*}/2\beta_2, \text{ conj.} \tag{4.11}$$

From (4.10), (4.11), we find the important relation

$$\partial_{\lambda^*} q_\mu = \partial_\mu q_{\lambda^*}, \text{ conj.} \tag{4.12}$$

Upon substituting (4.11), (4.12) into (4.8), we see that

$$'C^{\cdots\beta}_{\alpha\mu^*\nu} = C^{\cdots\beta}_{\alpha\mu^*\nu}, \text{ conj.} \tag{4.13}$$

Hence our theorem is proved.

We may also prove the converse theorem:

THEOREM 5. *If two unitary spaces are conformal and the affinor $C^{\cdots\beta}_{\alpha\mu^*\nu}$ is a conformal invariant, then the spaces are k-spaces.*

From (4.8), we find that the condition (4.13) is equivalent to the two conditions

$$\partial_{\lambda^*} q_\mu = \partial_\mu q_{\lambda^*}, \text{ conj.}, \tag{4.14}$$

$$\partial_\mu q_{\lambda^*} = \rho\, 'a_{\mu\lambda^*}, \text{ conj.} \tag{4.15}$$

From (3.8), we find

$$\partial_\mu \partial_{\lambda^*} \log \sigma = 2\rho\, 'a_{\mu\lambda^*}, \text{ conj.}, \tag{4.16}$$

where σ is real. Hence our theorem is proved.

UNIVERSITY OF TEXAS,
AUSTIN, TEX.

ON A GENERALIZED FORM OF PLATEAU'S PROBLEM

BY

R. COURANT

1. Introduction. In the variational problem of Plateau and in its generalization by Douglas, minimal surfaces S of least area are sought whose boundary Γ shall consist of monotonically described curves in euclidean space with coordinates x_1, x_2, x_3[1]. These surfaces are represented parametrically by position vectors $\mathfrak{x}(u, v)$ in a parameter domain B of the u, v-plane with the boundary C, in such a way that $\mathfrak{x}$ is continuous in $B+C$ and maps C monotonically on Γ. For the formulation of the problem and for the usual existence proof the monotonic character of the mapping of C on Γ is essential[2] (and implies as a consequence, incidentally, that C and Γ correspond to each other in a one-to-one way). However, from a geometrical viewpoint the restriction imposed on the parametric representation of the boundary is not natural. Therefore we replace the Plateau problem, here called *Problem* I, by another problem differing from the former by considerably weaker boundary conditions which do not stipulate continuity and monotonicity of the boundary mapping. This new "*Problem* II," to be formulated presently, is solved by the method developed in connection with "free boundaries"[3]. It is the purpose of the present note to prove that the solutions of the Plateau Problem I also solve the new Problem II; in other words, our widening of the range of competition does not affect the value of the minimum area. That this result is not obvious appears from the example at the end of this paper. We confine our proof to the case of a simply-connected surface S whose parameter-do-

---

Presented to the Society, April 27, 1940; received by the editors May 22, 1940.

(1) The restriction to three dimensions is not essential; the reasoning applies equally well in n dimensions.

(2) References—

[1] J. Douglas, *Minimal surfaces of higher topological structure*, Annals of Mathematics, (2), vol. 40 (1939), p. 205 ff.

[2] R. Courant, *The existence of minimal surfaces of given topological structure under prescribed boundary conditions*, Acta Mathematica, vol. 72 (1940), p. 51 ff.

[3] R. Courant, *The existence of a minimal surface of least area bounded by prescribed Jordan arcs and prescribed surfaces*, Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 97–101.

[4] R. Courant and N. Davids, *Minimal surfaces spanning closed manifolds*, Proceedings of the National Academy of Sciences, vol. 26 (1940), pp. 194–199.

[5] R. Courant, *Plateau's problem and Dirichlet's principle*, Annals of Mathematics, (2), vol. 38 (1937), p. 679 ff.

[6] R. Courant, *Conformal mapping of multiply connected domains*, Duke Mathematical Journal, vol. 5 (1939), p. 814 ff.

(3) See [2], [3], [4].

main B is the unit circle. The generalization to topologically higher cases does not present any new difficulty.

2. **Problem II. Generalized minimizing sequences.** In the unit circle B with the circumference C we consider "admissible vectors" $\mathfrak{x}(u, v)$, or in polar coordinates $\mathfrak{x}(r, \theta)$, defined by the conditions: $\mathfrak{x}$ is continuous and has piecewise-continuous(4) first derivatives in B; the boundary of $\mathfrak{x}$ is on the prescribed Jordan curve Γ in the x-space—this means that all limit points of $\mathfrak{x}(u, v)$ are on Γ as (u, v) tends to C; finally, instead of a monotonic mapping of C on Γ, we require that the curve $\mathfrak{x}(r, \theta)$ for fixed r shall be continuously deformable into the simply described curve Γ in such a way that for r sufficiently near to 1 this deformation can be effected within an arbitrarily small neighborhood of Γ. This topological condition is equivalent to the condition that, if H is any fixed simple polygon which is simply linked with Γ, the curve $\mathfrak{x}(r, \theta)$ shall, for r sufficiently near to 1, be simply linked with H(5).

Problem II requires constructing an admissible vector for which the Dirichlet integral

$$D(\mathfrak{x}) = \frac{1}{2}\int\int_B (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)dudv = \frac{1}{2}\int_0^{2\pi}\int_0^1 \left(\mathfrak{x}_r^2 + \frac{1}{r^2}\mathfrak{x}_\theta^2\right) r dr d\theta$$

attains a minimum value d. The existence of admissible vectors with finite $D(\mathfrak{x})$ is assumed, and certainly assured if Γ is rectifiable(6).

Problem I differs from Problem II by the restricting assumption that $\mathfrak{x}$ is continuous in $B+C$ and maps C monotonically on Γ. If k is the minimum value of $D(\mathfrak{x})$ in Problem I we certainly have $k \geqq d$, since vectors admissible in I are admissible in II. Our goal is to prove $d = k$.

In [2] and [4] it is shown that Problem II is solved by a minimal surface S. This solution is obtained as the limit of a sequence of vectors $\mathfrak{x}_n$ whose members satisfy the conditions of admissibility except that the boundary of $\mathfrak{x}_n$ need not be exactly on Γ but may be on a manifold Γ_n which tends to Γ as n increases, such that the maximum distance of points of Γ_n to Γ tends to zero. Sequences of this type are called *"admissible sequences."* The smallest lower limit of $D(\mathfrak{x}_n)$ for all such sequences is, as seen in [4], the same value d as for minimizing sequences of strictly admissible vectors. An admissible sequence $\mathfrak{x}_n$ for which $D(\mathfrak{x}_n) \to d$ is called *a generalized minimizing sequence.*

3. **Proof of the main theorem.** For the proof we shall apply a general deformation method(7) and a theorem on conformal mapping(8). To avoid non-

---

(4) This means that in any closed subdomain the continuity may be interrupted at a finite number of points or arcs with continuously turning tangent.

(5) As to the essential character of such linking conditions, compare [2], [4].

(6) See, e.g., [2], p. 54, Footnote 2.

(7) See [5], pp. 695, 707, and [6], p. 820.

(8) Also applied in [2] and [5].

essential technical complications we make the assumption that Γ is piecewise smooth, i.e., consists of a finite number of arcs having tangents continuous including the end-points.

In [2] and [4] it is shown that every generalized minimizing sequence $\mathfrak{x}_n$ can be replaced by such a sequence of harmonic vectors, which has a subsequence converging to a solution of II. The parameters u, v of each vector may be so chosen that the origin always corresponds to a point on the polygon H. In [2] and in the previous literature the following theorem on Problem I is established: if Γ_n is a sequence of simple polygons converging to Γ in the Fréchet sense, and if $\mathfrak{z}_n$ is admissible in Problem I for Γ_n, then the solution $\mathfrak{z}$ of I for Γ satisfies the inequality

$$D(\mathfrak{z}) \leqq \liminf D(\mathfrak{z}_n).$$

The goal of this paper is therefore attained if we construct a generalized minimizing sequence for II consisting of such vectors $\mathfrak{z}_n(u, v)$ which map C continuously and monotonically on Γ_n, and hence are admissible in the variational Problem I for the contour Γ_n.

The construction of $\mathfrak{z}_n$ will proceed in two steps. If the harmonic vector $\mathfrak{x}(u, v)$, or in polar coordinates $\mathfrak{x}(r, \theta)$, is a solution of Problem II with $D(\mathfrak{x}) = d$ and if $r_n < 1$ is a sequence of radii tending to 1, then $\mathfrak{x}_n(r, \theta) = \mathfrak{x}(r/r_n, \theta)$ is certainly a generalized minimizing sequence; the boundary $\mathfrak{x}_n(1, \theta)$ is an analytic curve Λ_n which, for given small $\epsilon = \epsilon_n$ and sufficiently large r_n, will be everywhere nearer to Γ then $\epsilon/2$.

We inscribe in Γ a polygon Γ_n whose edges have a length between $\epsilon^{1/2}$ and $2\epsilon^{1/2}$. By choosing the vertices of Γ_n on Γ and by choosing a suitably large unit length in the x-space we are sure that for sufficiently small ϵ the distance from every point of Γ_n to Γ is less than $\epsilon/2$, therefore to Λ_n less than ϵ. (Here the piecewise smoothness of Γ is used to ensure that the distance between Γ and Γ_n is small of a higher order than the edges of Γ_n.) For $\epsilon = \epsilon_n$ we choose a sequence tending to zero, so that Γ_n tends to Γ in the Fréchet sense; r_n, $\mathfrak{x}_n$, Λ_n are defined accordingly.

Now the first step of the construction consists in replacing the sequence $\mathfrak{x}_n$ by another generalized minimizing sequence $\mathfrak{y}_n(u, v) = \mathfrak{y}_n(r, \theta)$ such that $\mathfrak{y}_n$ is continuous in $B+C$ and maps C on the polygon Γ_n instead of on Λ_n. $\mathfrak{y}_n(1, \theta)$ is piecewise analytic in θ but not necessarily monotonic. The second step consists in showing that the sequence $\mathfrak{y}_n$ may be replaced by another generalized minimizing sequence $\mathfrak{z}_n$ with the same properties as $\mathfrak{y}_n$, but mapping C monotonically on Γ_n. Our main theorem will then be proved.

We start with the second step, assuming that the sequence $\mathfrak{y}_n$ has already been constructed. If $\mathfrak{y} = \mathfrak{y}_n(r, \theta)$ has piecewise analytic boundary values $\mathfrak{y}(1, \theta)$ which describe the polygon Γ_n not monotonically, there are but a finite number of "stationary" θ-intervals along which $\mathfrak{y}$ is a constant and a finite number of "folds," i.e., parts of Γ_n which are described more than

once and in different directions. We eliminate the stationary intervals and folds successively by conformal mapping. If A_0A_1: $\theta_0 \leqq \theta \leqq \theta_1$, is a stationary interval s on C, we map the circle B conformally on another unit circle B' with a slit s' so that the slit s corresponds to the arc s. This mapping transforms $\mathfrak{y}$ in B into $\mathfrak{y}'$ in $B'-s'$ and, because of the invariance of the Dirichlet integral, we have

$$D_{B'-s'}(\mathfrak{y}') = D_B(\mathfrak{y}).$$

But, since $\mathfrak{y}'$ is constant on s it is continuous in B' including its boundary. Denoting B' and $\mathfrak{y}'$ again by B and $\mathfrak{y}$, we have replaced $\mathfrak{y}=\mathfrak{y}_n$ by another vector with the same properties and at the same time reduced the number of stationary intervals by one. Repeating this process we eliminate in $\mathfrak{y}$ all stationary intervals.

The folds are treated in a similar way. Suppose the interval s: A_0PA_1 on C corresponds to a fold such that the arcs A_0P and PA_1 are mapped analytically on the same portion of an edge of Γ_n, each arc monotonically but the two arcs with different sense of description. Then those points on A_0P and PA_1 which have the same image point on Γ_n correspond to each other by an analytic transformation of θ. Hence, according to a procedure used in previous publications(9) (*sewing theorem*) we can identify such corresponding points by conformal mapping. Namely, we can map B on a unit circle B' with an analytic slit s' so that the interiors of the arcs A_0P and PA_1 correspond to the two edges of s' and that corresponding points on s go into opposite points on the two edges of s'. The end-points A_0, A_1, P are transformed into points or possibly into analytic slits. Again, $\mathfrak{y}$ is transformed into $\mathfrak{y}'$ in $B'-s'$ and the reasoning proceeds as above, since $\mathfrak{y}'$ remains continuous in B' after erasing s'. In this way, by a finite number of steps, all the folds are eliminated without any change of the Dirichlet integral. We therefore obtain, starting from $\mathfrak{y}_n$ a new generalized minimizing sequence $\mathfrak{z}_n$ as stipulated.

The remaining step, i.e., the construction of $\mathfrak{y}_n$ from $\mathfrak{x}_n$, is done by a general variational method(10). We consider the influence of certain deformations of the whole x-space on the variational integral. In our case these deformations are to affect only the neighborhood of Γ. They deform the analytic curve Λ_n given by $\mathfrak{x}(r_n, \theta)$ into the polygon Γ_n. Considering those points in the x-space which are images under $\mathfrak{x}_n(u, v)$ of points in B, we transform $\mathfrak{x}_n(u, v)$ into a vector $\mathfrak{y}_n(u, v)$, and we shall see to it that $\mathfrak{y}_n$ is again a generalized minimizing sequence. Our transformations will be described as composed of a finite number of "*pinching processes*":

(a) First, "δ-pinching around a point P" is defined as follows: assuming P to be the origin $x_1=x_2=x_3=0$, we define, for a given small δ and $t=t(\delta)>1$,

(9) See [2], p. 57 ff.
(10) See [2], p. 81.

the vector

$$y_i = p(r)x_i, \qquad i = 1, 2, 3,$$

where $r^2 = \sum x_i^2$ and

$$p = 1 \quad \text{for} \quad r > t\delta; \qquad p = 0 \quad \text{for} \quad r < \delta;$$

$$p = \sigma \log \frac{r}{\delta} \quad \text{for} \quad \delta < r < t\delta$$

with $\sigma = 1/\log t$. In other words, the deformation pinches the δ-sphere around P into P, leaves points outside of the larger sphere with the radius $t\delta$ unchanged, and moves every point on a straight line towards P. Substituting for x_i functions of u and v we obtain y_i as functions of u and v and thus a vector $\mathfrak{y}(u, v)$.

To compare the Dirichlet integral of $\mathfrak{y}$ with that of $\mathfrak{x}$ we consider in B the subset B^* where $p<1$, i.e., those points u, v whose images under $\mathfrak{x}(u, v)$ are nearer to P than $t\delta$. Let $D^* = D_{B^*}$ denote the Dirichlet integral over B^*. We then have, since $|\mathfrak{x}| = r$, $|x^i| \leqq r$, $|p| \leqq 1$[11],

$$\begin{aligned} D^*(\mathfrak{y}) &\leqq \frac{1}{2}\iint_{B^*} \sum_{i=1}^{3}\left[\left\{|px_u^i| + |\sigma x^i|\frac{1}{r^2}|\mathfrak{x}\mathfrak{x}_u|\right\}^2\right. \\ &\quad \left. + \left\{|px_v^i| + |\sigma x^i|\frac{1}{r^2}|\mathfrak{x}\mathfrak{x}_v|\right\}^2\right] du\,dv \\ &\leqq \frac{1}{2}\iint \sum[\{|x_u^i| + \sigma|\mathfrak{x}_u|\}^2 + \{|x_v^i| + \sigma|\mathfrak{x}_v|\}^2]\,du\,dv \\ &\leqq \frac{1}{2}\iint (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)(1 + 3\sigma^2 + 4\sigma)\,du\,dv; \end{aligned}$$

hence

$$(1) \qquad D^*(\mathfrak{y}) \leqq \frac{1}{2}\iint_{B^*} (1 + 3\sigma)^2(\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\,du\,dv = (1 + 3\sigma)^2 D^*(\mathfrak{x})$$

which shows that for small σ the influence of the transformation on the Dirichlet integral is small. We may write

$$(2) \qquad D(\mathfrak{y}) \leqq \frac{1}{2}\iint_{B} (1 + \eta)^2(\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\,du\,dv$$

where $\eta = 0$ outside B^* and

$$(3) \qquad |\eta| \leqq 3\sigma(\delta)$$

elsewhere.

[11] The coordinate indices are written as superscripts here and immediately below.

(b) Secondly, we introduce "δ-pinching around a straight segment," for which, without restriction of generality, we may choose

$$s: \quad x_2 = x_3 = 0; \qquad |x_1| < a + \alpha;$$

$t = t(\delta)$ and σ are defined as above and we assume

$$|a| > t\delta, \qquad |\alpha| > t\delta.$$

We define the transformation for $|x_1| < a$ by

$$y_1 = x_1, \qquad y_2 = p(\rho)x_2, \qquad y_3 = p(\rho)x_3,$$

with $\rho^2 = x_2^2 + x_3^2$ and $p(\rho)$ as above. For $x_1 > a$ and $x_2 < -a$ the transformation is defined as δ-pinching around the point $x_1 = a$ and $x_1 = -a$ on the segment s, respectively. By this pinching an essentially cylindrical bar around s with two half-spherical ends is contracted into s.

Again a simple appraisal shows that, for the subdomain B^* of B to which vectors $\mathfrak{x}$ affected by our pinching correspond, we have the same inequality (1) or (2) as in the pinching around a point.

(c) By carrying out successively N pinchings around different points and lines, possibly with different values of $\delta = \delta_i$, we obtain from $\mathfrak{x}$ a vector $\mathfrak{y}$ with

$$D(\mathfrak{y}) \leqq \frac{1}{2}\int\int_B (1 + \eta_1)^2 \cdots (1 + \eta_N)^2 (\mathfrak{x}_u^2 + \mathfrak{x}_v^2)\,du\,dv$$

where the η_i are defined as above. If no more than two quantities η_i differ from zero at the same point, we have

$$D(\mathfrak{y}) \leqq (1 + 3\sigma)^4 D(\mathfrak{x})$$

where σ now denotes the largest value $1/\log t_\nu$ employed in our pinchings.

(d) In applying the preceding results we consider, as above, a sequence of small values $\epsilon = \epsilon_n$ tending to zero, and corresponding polygons Γ_n and analytic curves Λ_n representing the boundary of $\mathfrak{x}_n$. We transform, for sufficiently large n and small $\epsilon = \epsilon_n$, $\mathfrak{x}_n$ into another vector $\mathfrak{y}_n$ as follows: with

$$\tau = \epsilon^{1/8}, \qquad \delta = \tau^6, \qquad t = \tau^{-1}, \qquad \sigma = \frac{8}{\log \epsilon^{-1}},$$

we perform a δ-pinching around each vertex of Γ_n. For small values of ϵ these deformations do not interfere with each other and transform Γ_n into itself, since the edges of Γ_n are all longer than $\tau^4 = \epsilon^{1/2}$ and the angles between adjacent edges remain above a fixed positive bound. We then subject each edge of Γ_n, except the ends, already pinched into the vertices, to a δ-pinching as under (b). In this deformation we choose as s a symmetric part of the edge by cutting off from each end the length $\frac{1}{2}\tau^6$ so that, if $l \geqq \tau^4$ is the length of the edge, we have

$$2a = l - \tau^6, \qquad \delta = \tau^3, \qquad t = \tau^{-1}, \qquad \delta t = \tau^{-7}.$$

Here the dimensions are again so balanced that for sufficiently small ϵ no interference with another edge occurs. Moreover it is obvious that the succession of all our pinchings results in a piecewise analytic deformation of $\mathfrak{x}_n$ into a vector $\mathfrak{y}_n$ such that the analytic curve Λ_n is transformed into Γ_n and that the condition for the inequality (1) is satisfied for $\mathfrak{x} = \mathfrak{x}_n$, $\mathfrak{y} = \mathfrak{y}_n$, $\sigma = \sigma_n = 8/\log \epsilon_n^{-1}$. Hence the $\mathfrak{y}$ are again a generalized minimizing sequence. The proof of our main theorem is thereby completed.

4. **Remarks.** While every solution of I is proved to be a solution of II, the converse is not necessarily true. The following example demonstrates that

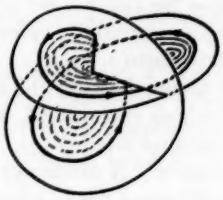

FIG. 1

for certain prescribed Jordan contours Γ the absolute minimum in the generalized problem can be attained by a vector $\mathfrak{x}(u, v)$ which maps the circle C in a continuous but non-monotonic way on Γ, such that on a part of Γ a fold is formed.

We consider a simply-connected minimal surface S with a branch point similar to that of the Riemann surface of $w^{1/2}$ and furnishing the absolute minimum for a Plateau Problem I([12]). Any continuous, non-self-intersecting curve Γ on S obviously bounds a simply connected minimal surface T on S

([12]) Apparently no instance of a minimal surface with branchpoint has been pointed out explicitly in the literature. However, the following procedure provides infinitely many examples: We start with a triangle OAB where OA and OB are straight segments of equal length forming an angle of 120 degrees and AB an arc (e.g., circular or polygonal) not in a plane with the points O, A, B. This triangle bounds a minimal surface M' of least area. Now, by reflecting M' successively on its straight edges issuing from O we obtain after 5 reflections a minimal surface M of the desired type with O as branchpoint and with *three* straight lines of self-intersection at angles of 120 degrees. That any simply connected portion S of M including the branchpoint O and having a piecewise smooth boundary curve furnishes an absolute minimum for the area, is inferred, for sufficiently flat triangles OAB from the classical theory. Our figure shows only the one line of self-intersection which is used for the construction.

Dr. C. L. Siegel in Princeton called the author's attention to a paper by B. Stessmann, *Periodische Minimalflächen*, Mathematische Zeitschrift, vol. 38 (1934), pp. 417–442, in which the Example 4 leads to a surface of the type described above. According to a personal communication, Professor Radó, using the principle of reflection, likewise has constructed an example of a simply connected minimal surface with branchpoint, bounded by a knotted clover-leaf curve. (This footnote was added January 15, 1941.)

solving Problem I for the boundary Γ. Figure 1 represents such a surface S topologically but with some metric exaggeration; for the sake of clarity, the curve Γ is indicated by a heavy line. The line Γ goes through the branch point and follows along a segment of a line of self-intersection c of S. T is represented separately in Figure 2. The portion T on S to the right of Γ has the

FIG. 2

minimum area k for Problem I. On the other hand we may, instead, cut the surface T along the line of intersection c as far as c coincides with Γ. T then becomes another simply connected domain T' with the same area and the same boundary Γ, except that the part of Γ along this cut now is described once in one direction and twice in the other, as indicated in the Figure 3. T is therefore a solution of II which is not a solution of I.

FIG. 3

NEW YORK UNIVERSITY
NEW YORK, N. Y.

PENCILS ON AN ALGEBRAIC VARIETY AND A NEW PROOF OF A THEOREM OF BERTINI

BY

OSCAR ZARISKI(1)

Introduction. A well known theorem of Bertini-Enriques on *reducible* linear systems of V_{r-1}'s on an algebraic V_r (i.e., linear systems in which each element is a reducible V_{r-1}) states that any such system, if free from fixed components, is composite with a pencil. The usual geometric proof of this theorem is based on another theorem of Bertini, to the effect that the *general* V_{r-1} of a linear system cannot have multiple points outside the singular locus of the variety V_r and the base locus of the system. This geometric proof has been subsequently completed and presented by van der Waerden under an algebraic form [3].

In this paper we give a new proof of the theorem of Bertini on reducible linear systems and we also extend this theorem to *irrational* pencils, i.e., pencils of genus $p>0$. Our proof does not make use of the second theorem of Bertini just quoted. In the case of pencils (linear or irrational), we first observe that a pencil $\{W\}$ on V_r is defined by a field P of algebraic functions of one variable which is a subfield of the field Σ of rational functions on V_r. The whole proof is then essentially based on the simple remark that *the pencil* $\{W\}$ *is composite with another pencil* $\{\overline{W}\}$, *defined by a field* $\overline{P}$, *if and only if* P *is a subfield of* $\overline{P}$. This property is a straightforward consequence of the geometric definition of composite pencils. As a matter of fact we prefer to define composite pencils by this property. At any rate, it is then true that *a pencil* $\{W\}$ *is non-composite if and only if the corresponding field* P *is maximally algebraic in* Σ. In the light of this approach to the question, the theorem of Bertini on reducible pencils is almost a direct consequence of the well known fact that an irreducible algebraic variety V, over a ground field K, is absolutely irreducible if K is maximally algebraic in the field of rational functions on V.

In the case of linear systems of dimension >1 the proof is even simpler, provided use is made of a certain lemma (Lemma 5). This lemma is, however, of interest in itself.

A sizable portion of the paper (Part I) is devoted to the development of the concept of a pencil and of the basic properties of pencils in the abstract case of an arbitrary ground field (of characteristic zero).

Presented to the Society, January 1, 1941; received by the editors July 1, 1940.

(1) Guggenheim Fellow.

I. Pencils of V_{r-1}'s on a V_r

1. Divisors of the first kind. Let V be an irreducible algebraic r-dimensional variety in a projective n-space, over an arbitrary ground field K of characteristic zero. We shall assume that V is *normal* in the projective space(2). Let Σ be the field of rational functions on V. Since V is normal, any irreducible $(r-1)$-dimensional subvariety Γ of V defines a prime divisor $\mathfrak{P}$ of Σ, i.e., an homomorphic mapping of Σ upon (Σ_1, ∞), where Σ_1 is the field of rational functions on Γ. There is also an associated $(r-1)$-dimensional valuation B of Σ/K, whose valuation ring $\mathfrak{L}$ is the quotient ring $Q(\Gamma)$ of Γ(3). It is well known that B (being of dimension $r-1$) is a *discrete valuation of rank* 1, i.e., its value group is the group of integers. If η is an element of Σ and if its value $v_B(\eta)$ in the valuation B is α (α a positive, negative, or zero integer), we shall say that η has *order* α *at the prime divisor* $\mathfrak{P}$, or *along the variety* Γ. We shall also say that η *vanishes to order* α *at* $\mathfrak{P}$, *or along* Γ, if $\alpha>0$, and that η *is infinite to the order* $-\alpha$ *at* $\mathfrak{P}$, *or along* Γ, if $\alpha<0$.

The prime divisors $\mathfrak{P}$ defined as above, by means of irreducible subvarieties of V *of dimension* $r-1$, shall be referred to as *divisors of the first kind* (*with respect to* V(4)). Dealing with the given normal variety V, we shall only deal with prime divisors of the first kind with respect to V. Concerning these we state the following well known theorem (Krull [1, p. 137, Vollständigkeitseigenschaft]):

---

(2) For the definition of normal varieties see our paper [4, p. 279, 283]. Our assumption implies that V is normal in the affine space, *for any choice of the hyperplane at infinity*. It is this weaker condition that really matters in our present treatment, rather than the condition that V be normal in the projective space. The restriction to normal varieties (either in the projective or in the above affine sense) is a sound principle from the standpoint of birational geometry. We have proved, in fact, that normal varieties exist in every class of birationally equivalent varieties. We have also associated with any given variety V a definite class of projectively related normal varieties, the *derived normal varieties of* V [4, p. 292], and therefore results proved for these can be readily restated as results concerning the original V. Finally, we point out that the class of varieties which are normal in the above affine sense includes the class of varieties which are free from singularities (in the projective space).

(3) We choose as hyperplane at infinity any hyperplane which does not contain Γ. Let $\xi_1, \cdots, \xi_n$ be the nonhomogeneous coordinates of the general point of V. Since the subvariety Γ is not at infinity, it is given by a prime ideal $\mathfrak{p}$ of the ring $\mathfrak{o}=K[\xi_1, \cdots, \xi_n]$. This ideal is $(r-1)$-dimensional and is minimal in $\mathfrak{o}$. Since V is normal, the ring $\mathfrak{o}$ is integrally closed in its quotient field Σ. Therefore the *quotient ring* $\mathfrak{o}_\mathfrak{p}$ $(=Q(\Gamma))$ is a valuation ring $\mathfrak{L}$. The residue field of the corresponding valuation B is the quotient field of the residue class ring $\mathfrak{o}/\mathfrak{p}$, and hence coincides with Σ_1. The homomorphic mapping of $\mathfrak{L}$ upon Σ_1 defines the prime divisor $\mathfrak{P}$.

(4) The notion of a prime divisor *of the second kind with respect to* V is defined as follows. Any prime divisor $\mathfrak{P}$ of Σ is by definition a homomorphic mapping of Σ upon (Σ_1, ∞), where Σ_1 is a field of algebraic functions of $r-1$ variables. By a proper choice of the hyperplane at infinity we may arrange matters so that none of the coordinates ξ_i is mapped upon the symbol ∞. The elements of the ring $\mathfrak{o}$ (see Footnote 3) which are mapped upon the zero element of Σ form then a prime ideal $\mathfrak{p}$ in $\mathfrak{o}$, and this prime ideal defines an irreducible subvariety Γ of V. If Γ is of dimension $r-1$, then our divisor $\mathfrak{P}$ is of the first kind and is uniquely determined by Γ. If Γ is

If an element η of Σ is transcendental over K, *then the set of prime divisors of the first kind along which η is infinite is finite and non-empty.*

A *divisor of the first kind* will be by definition a power product $\mathfrak{A}=\mathfrak{P}_1^{a_1}\cdots\mathfrak{P}_h^{a_h}$, where $\mathfrak{P}_1,\cdots,\mathfrak{P}_h$ are *prime* divisors of the first kind and where $\alpha_1,\cdots,\alpha_h$ are positive, negative, or zero integers. If all the α are positive integers or zero, then $\mathfrak{A}$ is said to be an *integral divisor.*

$\mathfrak{A}$ is said to be a *principal divisor* if there exists an element η in the field Σ such that η has order α_i at $\mathfrak{P}_i$ $(i=1, 2,\cdots, h)$ and order 0 at any other prime divisor of the first kind. Notation: $\mathfrak{A}=(\eta)$.

2. **Definition of a pencil.** Let P be a subfield of Σ *containing the ground field* K and being of *degree of transcendency* 1 *over* K. By means of such a subfield P we proceed to define a collection $\{W\}$ of $(r-1)$-dimensional subvarieties W of V, and namely *one* W for each *place*, or *prime divisor*, of P/K.

Given an $(r-1)$-dimensional irreducible subvariety Γ of V, let $\mathfrak{P}$ be the corresponding prime divisor of Σ/K and let B be the valuation defined by $\mathfrak{P}$. The valuation B induces a valuation B_1 in the field P. The valuation B_1 is either the trivial valuation, in which every element of P (different from 0) has value zero, or B_1 is a non-trivial valuation. In the first case, the mapping of P in the divisor $\mathfrak{P}$ is an isomorphism. In the second case, P is mapped by $\mathfrak{P}$ homorphically upon a field which is algebraic over K. This mapping defines a *place*, or a prime divisor $\mathfrak{p}$ of P/K; it is the prime divisor which is also directly defined by the non-trivial valuation B_1. We say in this second case that the irreducible $(r-1)$-dimensional subvariety Γ, or the corresponding prime divisor $\mathfrak{P}$, *corresponds to the place* $\mathfrak{p}$ *of* P/K.

It is not difficult to see that there is at most a *finite number of irreducible V_{r-1}'s which correspond to a given place* $\mathfrak{p}$ *of* P/K. In fact, let t be an element of P whose order at the place $\mathfrak{p}$ is *positive*. Then it is clear that the prime divisors of the first kind which correspond to $\mathfrak{p}$ must be among those prime factors of the principal divisor (t) whose exponents are positive. The number of such prime factors is finite, since, by a previously stated theorem, $1/t$ is infinite only at a finite number of prime divisors of the first kind.

We shall see later that *to each place* $\mathfrak{p}$ *there corresponds at least one irreducible V_{r-1}* on V.

Let $\Gamma_1,\cdots,\Gamma_m$ be the irreducible V_{r-1}'s on V which correspond to the

---

of dimension *less* than $r-1$, then $\mathfrak{P}$ is of the second kind with respect to V. There exist infinitely many prime divisors $\mathfrak{P}$ of the second kind leading to one and the same irreducible subvariety Γ of dimension $<r-1$.

It is well known that a prime divisor $\mathfrak{P}$ of Σ/K defines a discrete valuation B of Σ, of rank 1. The valuation ring $\mathfrak{L}$ of B is the set of all elements of Σ which are mapped upon elements of Σ_1. Our condition on the choice of the hyperplane at infinity implies that $\mathfrak{o}\subset\mathfrak{L}$. The subvariety Γ is called *the center* of the valuation B on V. A prime divisor $\mathfrak{P}$ of Σ/K is of the first kind with respect to V if and only if the center of the valuation defined by $\mathfrak{P}$ is a subvariety of V of *dimension* $r-1$ (and not less).

given place $\mathfrak{p}$ of P/K. Let t be a uniformizing parameter at $\mathfrak{p}$. We attach to each Γ_i *a multiplicity* h_i: namely, h_i *shall be the order to which* t *vanishes along* Γ_i. It is clear that h_i is a positive integer and that it is independent of the choice of the uniformizing parameter t(5). We regard the variety

$$W_{\mathfrak{p}} = h_1\Gamma_1 + \cdots + h_m\Gamma_m$$

as *the total subvariety* of V which corresponds to the place $\mathfrak{p}$. We define *the pencil* $\{W\}$ as *the totality of all* $W_{\mathfrak{p}}$ *obtained as* $\mathfrak{p}$ *varies on the Riemann surface of* P/K.

3. **Adjunction of indeterminates.** In order to derive the basic properties of pencils, we proceed to give an explicit construction of the pencil $\{W\}$ based on Kronecker's method of indeterminates. Let $\xi_1, \cdots, \xi_n$ be, as before, the nonhomogeneous coordinates of the general point of V(6). We introduce the $r-1$ forms:

$$\eta_i = u_{i1}\xi_1 + \cdots + u_{in}\xi_n, \qquad i = 1, 2, \cdots, r-1,$$

where the $n(r-1)$ elements u_{ij} are indeterminates. These indeterminates we adjoin to the field Σ, getting a field $\Sigma^* = \Sigma(\{u_{ij}\}) = \Sigma(u_{11}, \cdots, u_{r-1,n})$. The field Σ^* is a pure transcendental extension of Σ, of degree of transcendency $n(r-1)$ over Σ. We also consider the fields:

$$\mathrm{P}^* = \mathrm{P}(\{u_{ij}\}, \{\eta_i\}),$$
$$\mathrm{K}^* = \mathrm{K}(\{u_{ij}\}, \{\eta_i\}).$$

The elements $\eta_1, \cdots, \eta_{r-1}$ *are algebraically independent over the field* $\mathrm{P}(\{u_{ij}\})$. For, if $\eta_1, \cdots, \eta_{r-1}$ and the u_{ij} satisfied an algebraic relation with coefficients in P, then by specializing the u_{ij}, $u_{ij} \to c_{ij} \in \mathrm{K}$, we could get an algebraic relation over P between any $r-1$ of the elements ξ_i(7). This is impossible, since Σ is of degree of transcendency $r-1$ over P.

From the algebraic independence of the η_i it follows that Σ^* and P^* have the same degree of transcendency over K, and that P^* is of degree of transcendency 1 over K^*. Hence:

(5) If t' is another uniformizing parameter at $\mathfrak{p}$, then t'/t has order zero at $\mathfrak{p}$, and consequently t'/t has also order zero along Γ_i.

(6) Generally speaking, the proper procedure would have been to use the homogeneous coordinates of the general point of V (Zariski [4, p. 284]), thus avoiding special considerations for *divisors of infinity* (i.e., prime divisors of the first kind, at which at least one of the elements ξ_i is infinite and which therefore correspond to the irreducible components of the section of V with the hyperplane at infinity), However, the use of homogeneous coordinates would have required introductory definitions and proofs concerning homogeneous prime divisors and similar concepts associated with homogeneous coordinates. The size of such an introduction would be out of proportion to the limited object of this paper. We therefore prefer to deal with nonhomogeneous coordinates, also because in the present case the special considerations for the divisors at infinity are very simple and brief.

(7) See van der Waerden [2, lemma on page 17].

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

Σ^* is an algebraic extension of P^*. If K^* is taken as ground field, then P^* is a field of algebraic functions of one variable.

Given a prime divisor $\mathfrak{p}$ of the field P/K, we shall want to extend $\mathfrak{p}$ to a prime divisor of the field P^*/K^*. This extension is based on the following lemma, which we shall use also later on in a different connection:

LEMMA 1. *Let $\Omega^* = \Omega(x_1, \cdots, x_m)$ be a pure transcendental extension of a field Ω (i.e., $x_1, \cdots, x_m$ are algebraically independent over Ω). Given a valuation B of Ω, there exists one and only one extended valuation B^* of Ω^* such that the B^*-residues of $x_1, \cdots, x_m$ are $\neq \infty$ and are algebraically independent over the residue field of B.*

Proof. Let B^* be a valuation of Ω^* satisfying the conditions of the lemma, and let $\eta^* = f(x_1, \cdots, x_m)/g(x_1, \cdots, x_m)$, $f, g \in \Omega[x_1, \cdots, x_m]$, be an arbitrary element of Ω^*. Some of the coefficients of the rational function f/g may have negative values in B. However, if we divide through f and g by a coefficient of minimum value, we get a rational function whose coefficients have *finite* B-residues, *not all zero*. We may assume then that the rational function f/g already satisfies this condition. Let us assume that the B-residues of the coefficients of the denominator g are all zero. Then the B^*-residue of $g(x_1, \cdots, x_m)$ is zero, since the B^*-residues of $x_1, \cdots, x_m$ are $\neq \infty$. On the other hand, the B-residues of the coefficients of the numerator f are necessarily not all zero. Hence the B^*-residue of f is different from zero, since the B^*-residues of $x_1, \cdots, x_m$ are algebraically independent over the residue field of B. Consequently the B^*-residue of η^* is ∞.

On the other hand, assuming that the B-residues of the coefficients of g are not all zero, we conclude in a similar fashion that the B^*-residue of g is $\neq 0$, while the B^*-residue of f is $\neq \infty$. Hence the B^*-residue of η^* is $\neq \infty$.

Hence *the valuation ring $\mathfrak{o}^*$ of B^* consists of all quotients f/g such that the B-residues of the coefficients are all finite and the B-residues of the coefficients of g are not all zero.* This shows that B^* is uniquely determined. On the other hand, the set of all such quotients is a ring $\mathfrak{o}^*$ satisfying the condition that if $\eta^* \notin \mathfrak{o}^*$, then $1/\eta^* \subset \mathfrak{o}^*$. Hence $\mathfrak{o}^*$ is a valuation ring. It is then immediately seen that the corresponding valuation B^* of Ω^* is an extension of B (i.e., $\mathfrak{o}^* \cap \Omega$ is the valuation ring of B) and satisfies the condition of the lemma.

We point out that *the value group of B^* is the same as the value group of B* and that *the residue field of B^* is a pure transcendental extension of the residue field of B, the adjoined transcendentals being the B^*-residues of $x_1, \cdots, x_m$.* The proofs of these assertions are straightforward. We also wish to point out explicitly that B^* *depends on the particular set of generators $x_1, \cdots, x_m$ of Ω^*/Ω.* Thus, the set of generators $cx_1, x_2, \cdots, x_m$, $c \subset \Omega$, defines an extended valuation of Ω^* which is different from B^*, *whenever $v_B(c) \neq 0$.*

We apply the above lemma to the fields P, P^*, of which the second is a

pure transcendental extension of P. As generators of P^*/P we take the elements u_{ij} and η_i. Let $\mathfrak{p}$ be a prime divisor of P/K, B_1 the corresponding valuation of P, B_1^* the extended valuation of P^*. If Δ_1 is the residue field of B_1, then $\Delta_1^* = \Delta_1(\{u_{ij}^*\}, \{\eta_i^*\})$ is the residue field of B_1^*, where the u_{ij}^* and η_i^* are the B_1^*-residues of the u_{ij} and η_i. The field Δ_1 contains a subfield simply isomorphic to K, which we may identify with K. The fields $K(\{u_{ij}^*\}, \{\eta_i^*\})$, $K(\{u_{ij}\}, \{\eta_i\})$ are simply isomorphic, in view of the algebraic independence of the residues u_{ij}^*, η_i^* with respect to Δ_1. Hence the valuation B_1^* defines a *prime divisor* $\mathfrak{p}^*$ of the field P^*/K^*, i.e., if K^* is taken as ground field and if P^* is regarded as a field of functions of one variable. *Hence we have associated with each prime divisor* $\mathfrak{p}$ *of* P/K *a prime divisor* $\mathfrak{p}^*$ of P^*/K^*:

$$\mathfrak{p} \to \mathfrak{p}^*. \tag{1}$$

Since B_1^* is an extension of B_1, it is clear that *distinct prime divisors of* P/K *extend to distinct prime divisors of* P^*/K^*.

We now consider the field Σ^*. It is an algebraic extension of P^*. Hence the prime divisor $\mathfrak{p}^*$ factors into a power product of prime divisors of Σ^*/K^*. Let

$$\mathfrak{p}^* = \mathfrak{P}_1^{*h_1} \cdots \mathfrak{P}_m^{*h_m}. \tag{2}$$

The following sections are devoted to the proof of the following assertions:

(A) *Each prime divisor* $\mathfrak{P}_i^*$ $(i = 1, 2, \cdots, m)$ *induces in* Σ *a prime divisor* $\mathfrak{P}_i$ *of the first kind.*

(B) *The prime divisors* $\mathfrak{P}_1, \cdots, \mathfrak{P}_m$ *are distinct.*

(C) *If* Γ_i *is the irreducible* V_{r-1} *defined on* V *by the prime divisor* $\mathfrak{P}_i$, *then* $h_1\Gamma_1 + \cdots + h_m\Gamma_m$ *is the member* $W_{\mathfrak{p}}$ *of the pencil* $\{W\}$ *which corresponds to the place* $\mathfrak{p}$ *of* P/K (in the sense of the definition given in §2).

4. **The induced prime divisors** $\mathfrak{P}_i$. We consider one of the prime divisors $\mathfrak{P}_i^*$ in (2), say $\mathfrak{P}_1^*$, and we denote by B^* the valuation of Σ^* defined by $\mathfrak{P}_1^*$. In Σ, a subfield of Σ^*, B^* induces a valuation B. Since Σ^* is of degree of transcendency $(r-1)n$ over Σ, *the residue field of* B^* *can be at most of degree of transcendency* $(r-1)n$ *over the residue field of* B(8). Since the residue field of B^* is an algebraic extension of K^*, it is of degree of transcendency $(r-1)n + (r-1)$

(8) The proof of this assertion is immediate. Let, quite generally, Σ^* be an extension field of a field Σ, of degree of transcendency ρ over Σ, and let B and B^* be respectively a valuation of Σ and an extended valuation of Σ^*. Let us assume that the residue field of B^* is of degree of transcendency $\geq \rho$ over the residue field of B. We consider ρ elements $\omega_1, \cdots, \omega_\rho$ of Σ^* whose B^*-residues $\omega_1^*, \cdots, \omega_\rho^*$ are algebraically independent over the residue field Δ of B. It is clear that $\omega_1, \cdots, \omega_\rho$ are algebraically independent over Σ. Hence, by the remark at the end of the proof of Lemma 1, it follows that the residue field of the valuation induced in the field $\Sigma(\omega_1, \cdots, \omega_\rho)$ by the valuation B^*, is $\Delta(\omega_1^*, \cdots, \omega_\rho^*)$. Now Σ^* is an algebraic extension of the field $\Sigma(\omega_1, \cdots, \omega_\rho)$. Hence the residue field of B^* is an algebraic extensions of $\Delta(\omega_1^*, \cdots, \omega_\rho^*)$, and therefore its degree of transcendency over Δ is exactly ρ.

over K. Hence the degree of transcendency of the residue field of B over K must be not less than $r-1$. On the other hand, $\mathfrak{P}_1^*$ induces in P the non-trivial prime divisor $\mathfrak{p}$. *Hence B, as a non-trivial valuation, is of dimension $r-1$.* It defines in Σ a prime divisor $\mathfrak{P}_1$. Thus *the prime divisors $\mathfrak{P}_1^*, \cdots, \mathfrak{P}_m^*$ induce in Σ/K prime divisors $\mathfrak{P}_1, \cdots, \mathfrak{P}_m$.*

We next show that the *prime divisors $\mathfrak{P}_1, \cdots, \mathfrak{P}_m$ are of the first kind with respect to V.* Let us consider for instance the divisor $\mathfrak{P}_1$. We examine separately two cases, according as the center of $\mathfrak{P}_1$ (i.e., of the corresponding valuation [see Footnote 4]) on V is or is not at finite distance.

First case. The center of $\mathfrak{P}_1$ is at finite distance (i.e., it does not lie in the hyperplane at infinity). In this case the B-residues $\xi_1^*, \cdots, \xi_n^*$ of $\xi_1, \cdots, \xi_n$ are all different from ∞. Let u_j^*, η_i^* be the B^*-residues of the elements u_{ij}, η_i. These B^*-residues are algebraically independent over K. On the other hand we have: $\eta_i^* = u_{i1}^*\xi_1^* + \cdots + u_{in}^*\xi_n^*$. Hence $r-1$ of the elements ξ_i^* must be algebraically independent over K. This shows that the center of $\mathfrak{P}_1$ on V is of dimension $r-1$, whence $\mathfrak{P}_1$ is of the first kind.

Second case. The center of $\mathfrak{P}_1$ is at infinity. In this case some of the elements ξ_i have negative values in B. Without loss of generality we may assume that $0 > v_B(\xi_1) = \min\ (v_B(\xi_1), \cdots, v_B(\xi_n))$. We consider the following projective transformation of coordinates:

$$\xi_1' = 1/\xi_1, \qquad \xi_i' = \xi_i/\xi_1, \qquad i = 2, \cdots, n. \tag{3}$$

With respect to the new nonhomogeneous coordinates ξ_i' the center of $\mathfrak{P}_1$ is at finite distance. To show that $\mathfrak{P}_1$ is of the first kind, we have to show that among the B-residues $\xi_1'^*, \xi_2'^*, \cdots, \xi_n'^*$ of the ξ_i' there are $r-1$ which are algebraically independent over K. To show this we observe that we have:

$$\eta_i = \frac{u_{i1} + u_{i2}\xi_2' + \cdots + u_{in}\xi_n'}{\xi_1'}. \tag{4}$$

Hence passing to the B^*-residues and noting that the B-residue of ξ_1' is zero we find:

$$u_{i1}^* + u_{i2}^*\xi_2'^* + \cdots + u_{in}^*\xi_n'^* = 0. \tag{5}$$

These relations show that the $r-1$ elements u_{ij}^* belong to the field $K(\{u_{ij}^*\}, j>1; \xi_2'^*, \cdots, \xi_n'^*)$. Since the $(r-1)n$ elements u_{ij}^* are algebraically independent over K, $r-1$ of the residues $\xi_2'^*, \cdots, \xi_n'^*$ must be algebraically independent over K, q.e.d.

Thus assertion (A) is fully established.

5. **Proof of** (B). Let us consider one of the divisors $\mathfrak{P}_1, \cdots, \mathfrak{P}_m$, say $\mathfrak{P}_1$, and let, as before, B be the corresponding valuation of Σ, while B^* is the valuation of Σ^* defined by $\mathfrak{P}_1^*$. We know that B^* is an extended valuation of B. We also know that Σ^* is a pure transcendental extension of Σ. By Lemma 1,

B^* is uniquely determined by B if a set of generators of Σ^* over Σ is known whose B^*-residues are algebraically independent over the residue field of B. Since $\mathfrak{P}_1^*, \cdots, \mathfrak{P}_m^*$ are distinct divisors, the assertion (B) will be proved if we show that *it is possible to exhibit such a set of generators which depend only on* B. Let us first suppose that the center of B is at finite distance. The residue field of B is the field $K(\xi_1^*, \cdots, \xi_n^*)$, where the ξ_i^* are the B-residues of the ξ_i. Since the field $K(\{u_{ij}^*\}, \{\eta_i^*\})$ is of degree of transcendency $(r-1)n+(r-1)$ over K and since this field is contained in the field $K(\{u_{ij}^*\}, \xi_1^*, \cdots, \xi_n^*)$, it follows that the $(r-1)n$ residues u_{ij}^* are algebraically independent over the field $K(\xi_1^*, \cdots, \xi_n^*)$ (since this last field is of degree of transcendency $r-1$ over K). Hence, *if the center of* B *is at finite distance, the* $(r-1)n$ *elements* u_{ij} *form a set of generators of the desired nature.*

Let us now assume that the center of B is at infinity, and let, say, $0>v_B(\xi_1)=\min\ (v_B(\xi_1), \cdots, v_B(\xi_n))$. In this case the B^*-residues u_{ij}^* of the u_{ij} are not algebraically independent over the residue field $K(\xi_2'^*, \cdots, \xi_n'^*)$ of B, in view of (5). However, the elements $u_{ij}, j>1$, and $\eta_1, \cdots, \eta_{r-1}$ also form a set of generators of Σ^* over Σ. We assert that the B^*-*residues of these elements are algebraically independent over* $K(\xi_2'^*, \cdots, \xi_n'^*)$. To show this, we observe that in view of the relations (5), the field

$$\Delta^* = K(\xi_2'^*, \cdots, \xi_n'^*; \{u_{ij}^*\}, j>1; \eta_1^*, \cdots, \eta_{r-1}^*)$$

contains the subfield $K(\{u_{ij}^*\}, \eta_1^*, \cdots, \eta_{r-1}^*)$. Since this last field is of degree of transcendency $(r-1)n+(r-1)$ over K, and since $K(\xi_2'^*, \cdots, \xi_n'^*)$ is of degree of transcendency $r-1$ over K, it follows that the $(n-1)r$ residues of u_{ij}^* $(j>1)$, $\eta_1^*, \cdots, \eta_{r-1}^*$, are algebraically independent over $K(\xi_2'^*, \cdots, \xi_n'^*)$, q.e.d.

The proof of assertion (B) is now complete.

6. **Proof of** (C). Let $\Gamma_1, \cdots, \Gamma_m$ be the $(r-1)$-dimensional irreducible subvarieties of V which correspond to the prime divisors $\mathfrak{P}_1, \cdots, \mathfrak{P}_m$. Let $W_\mathfrak{p}$ be the member of the pencil $\{W\}$ which corresponds to the place $\mathfrak{p}$ of P/K, in the sense of the definition given in §2. Since each prime divisor $\mathfrak{P}_i^*$, and hence also each $\mathfrak{P}_i$, induces in P/K the prime divisor $\mathfrak{p}$, *it follows that* $\Gamma_1, \cdots, \Gamma_m$ *are components of* $W_\mathfrak{p}$.

Let t be an element of P which has order 1 at $\mathfrak{p}$. Then t, considered as an element of P^*, also has order 1 at $\mathfrak{p}^*$. In view of (2), t has order h_i at $\mathfrak{P}_i^*$. Since the value group of the valuation of Σ defined by $\mathfrak{P}_i$ is the same as the value group of the valuation of Σ^* defined by $\mathfrak{P}_i^*$ (see remark at the end of the proof of Lemma 1), it follows that t *has also order* h_i *at* $\mathfrak{P}_i$. We conclude that Γ_i *is an* h_i*-fold component of* $W_\mathfrak{p}$ (in the sense of the definition of §2).

To complete the proof of (C), we have to show that if Γ is an irreducible component of $W_\mathfrak{p}$, then Γ coincides with one of the varieties $\Gamma_1, \cdots, \Gamma_m$. Let $\mathfrak{P}$ be the prime divisor of Σ defined by Γ and let B be the corresponding valua-

tion of Σ. Let B^* be *the* extended valuation of B in Σ^* defined by the following condition: (1) if Γ is at finite distance, then the B^*-residues of the $(r-1)n$ elements u_{ij} are algebraically independent over the residue field of B; (2) if $\mathfrak{P}$ is a divisor at infinity, and if, say, $0>v_B(\xi_1)=\min(v_B(\xi_1),\cdots,v_B(\xi_n))$, then the B^*-residues of the $(r-1)n$ elements $u_{ij}(j>1)$, $\eta_1,\cdots,\eta_{r-1}$, are algebraically independent over the residue field of B.

In view of Lemma 1, the valuation B^* is well defined in either case. We assert that the *B^*-residues of the elements u_{ij}, η_i are algebraically independent over* K.

Proof. *First case. $\mathfrak{P}$ is a divisor at finite distance.* Let u_{ij}^*, η_i^*, ξ_j^* denote B^*-residues of the corresponding elements u_{ij}, η_i, ξ_j. We have $\eta_i^*=u_{i1}^*\xi_1^*+\cdots+u_{in}^*\xi_n^*$. By our construction of B^*, the elements u_{ij}^* are algebraically independent over the field $K(\xi_1^*,\cdots,\xi_n^*)$. The rest of the proof is based on the van der Waerden lemma quoted in Footnote 7 and is identical with the proof that $\eta_1,\cdots,\eta_{r-1}$ are algebraically independent over $P(\{u_{ij}\})$.

Second case. $\mathfrak{P}$ is a divisor at infinity, and $0>v_B(\xi_1)=\min(v_B(\xi_1),\cdots,v_B(\xi_n))$. By our construction of B^*, the residues u_{ij}^*, $j>1$, are algebraically independent over the field $K(\xi_2'^*,\cdots,\xi_n'^*)$, where the ξ_i' are given by the relations (3)(9). By a specialization argument similar to the one used in the preceding case and based on van der Waerden's lemma, it follows that the $r-1$ forms $u_{i2}^*\xi_2'^*+\cdots+u_{in}^*\xi_n'^*$ are algebraically independent over the field $K(\{u_{ij}^*\}, j>1)$. But these forms are equal to the elements $-u_{i1}^*$, in view of the relations (5). Hence the $(r-1)n$ elements u^* are algebraically independent over K. On the other hand, the residues $\eta_1^*,\cdots,\eta_{r-1}^*$ are, by hypothesis, algebraically independent over the field $K(\xi_2'^*,\cdots,\xi_n'^*;\{u^*\}, j>1)$. Hence *a fortiori* these residues are algebraically independent over the subfield $K(\{u^*\})$, and this proves our assertion.

The fact that the B^*-residues u_{ij}^*, η_i^* are algebraically independent over K, implies that B^* is a valuation of Σ^* *over the ground field* K^*, i.e., *B^* defines a prime divisor $\mathfrak{P}^*$ of* Σ^*/K^*. This prime divisor $\mathfrak{P}^*$ lies over a definite prime divisor $^*\mathfrak{p}^*$ of P^*/K^*. To show that $\mathfrak{P}^*$ coincides with one of the prime divisors $\mathfrak{P}_1^*,\cdots,\mathfrak{P}_m^*$, it is only necessary to show that $^*\mathfrak{p}^*=\mathfrak{p}^*$. Now both $\mathfrak{p}^*$ and $^*\mathfrak{p}^*$ induce in the field P/K one and the same prime divisor, namely $\mathfrak{p}$, since, by hypothesis, Γ is a component of $W_{\mathfrak{p}}$. In view of Lemma 1, the relation $^*\mathfrak{p}^*=\mathfrak{p}^*$ will follow if we prove that the residues of the $(n+1)(r-1)$ elements u_{ij}, η_i in the divisor $^*\mathfrak{p}^*$ are algebraically independent over the residue field of the divisor $\mathfrak{p}$. But this is obvious, since this last residue field is algebraic over K and since we have just proved that the B^*-residues of the elements u_{ij}, η_i are algebraically independent over K. This completes the proof of assertion (C).

7. **Order of $W_{\mathfrak{p}}$.** Let $\eta_i=u_{i1}\xi_1+\cdots+u_{in}\xi_n$, $i=1,2,\cdots,r$, where the

(9) Note that since V is normal, $K(\xi_1'^*,\cdots,\xi_n'^*)$ is the residue field of B, and also that $\xi_1'^*=0$.

nr elements u_{ij} are indeterminates. We define the *order of* V as the relative degree $[\Sigma(\{u_{ij}\}):\mathrm{K}(\{u_{ij}\}, \eta_1, \cdots, \eta_r)]$. If we apply this definition to each irreducible component Γ_i of $W_\mathfrak{p}$, we can speak of the *order of* $W_\mathfrak{p}$: it shall be, by definition, the sum of the orders of $\Gamma_1, \cdots, \Gamma_m$, each counted to its multiplicity $h_1, \cdots, h_m$, respectively.

Given a field Ω of algebraic functions of one variable, over a ground field K, and given a prime divisor $\mathfrak{p}$ in Ω/K, by *the degree of* $\mathfrak{p}$ we mean the relative degree $[\Delta:\mathrm{K}]$, where Δ is the residue field of $\mathfrak{p}$.

From this definition it follows immediately that the prime divisor $\mathfrak{p}$ of P/K and the extended prime divisor $\mathfrak{p}^*$ of $\mathrm{P}^*/\mathrm{K}^*$ have the same degree, since the residue field of $\mathfrak{p}^*$ is a pure transcendental extension of the residue field of $\mathfrak{p}$ (see Lemma 1 and the application to the fields P, P*, §3). Let d be the common degree of $\mathfrak{p}$ and $\mathfrak{p}^*$. Let $d_1, \cdots, d_m$ be the degrees of the prime divisors $\mathfrak{P}_1^*, \cdots, \mathfrak{P}_m^*$ respectively. It is well known that the relation (2) implies

$$vd = h_1d_1 + \cdots + h_md_m, \tag{6}$$

where v *is the relative degree* $[\Sigma^*:\mathrm{P}^*]$. The relation (6) has a geometric interpretation, namely, that *the right-hand side* $h_1d_1+\cdots+h_{1n}d_{1n}$ *is the order of* $W_\mathfrak{p}$. To prove this, we have to show that d_i *is the order of* Γ_i.

Proof. *First case.* Γ_1 *is at finite distance.* The nonhomogeneous coordinates of the general point of Γ_1 are $\xi_1^*, \cdots, \xi_n^*$, where the ξ_i^* are the residues of the ξ_i in the valuation B defined by the prime divisor $\mathfrak{P}_1$. The field $\Delta=\mathrm{K}(\xi_1^*, \cdots, \xi_n^*)$ is the residue field of B. The residue field of the extended valuation B^* in Σ^* is (§5) $\mathrm{K}(\xi_1^*, \cdots, \xi_n^*, \{u_{ij}^*\})$. The degree of $\mathfrak{P}_1^*$ is the relative degree

$$d_1 = [\Delta(\{u_{ij}^*\}):\mathrm{K}(\{u_{ij}^*\}, \{\eta_i^*\})].$$

Since the u_{ij}^* are algebraically independent over Δ, the relation above shows that d_1 is also the order of Γ_1.

Second case. Γ_1 *is at infinity*, and $0>v_B(\xi_1)=\min(v_B(\xi_1), \cdots, v_B(\xi_n))$. In this case the general point of Γ_1 is $(\xi_1'^*, \xi_2'^*, \cdots, \xi_n'^*)$, where the ξ_i' are defined by the transformation (3). But since $\xi_1'^*=0$, Γ_1 actually lies in an S_{n-1} (in the hyperplane at infinity of S_n), and the general point of Γ_1 is $(\xi_2'^*, \cdots, \xi_n'^*)$. The field $\Delta=\mathrm{K}(\xi_2'^*, \cdots, \xi_n'^*)$ is the residue field of B. By §5, the residue field of B^* is $\Delta(\{u_{ij}^*\}, j>1; \{\eta_i^*\})$, and we have

$$d_1 = [\Delta(\{u_{ij}^*\}, j>1; \{\eta_i^*\}):\mathrm{K}(\{u_{ij}^*\}, \{\eta_i^*\})].$$

Since $\eta_1^*, \cdots, \eta_{r-1}^*$ are algebraically independent over the field $\Delta(\{u_{ij}^*\}, j>1)$, we can also write:

$$d_1 = [\Delta(\{u_{ij}^*\}, j>1):\mathrm{K}(\{u_{ij}^*\}, j>1; \{u_{i1}^*\})].$$

The assertion that d_1 is the order of Γ_1 now follows from the relations (5).

In view of (6), we may now state the following result: *the quotient*

$$\frac{\textit{order of } W_{\mathfrak{p}}}{\textit{degree of } \mathfrak{p}}$$

is the same for all members of the pencil $\{W\}$. In the particular case *when* K *is algebraically closed*, the degree of $\mathfrak{p}$ is always 1, and hence *all the members of the pencil* $\{W\}$ *have the same order.*

8. **The base points of the pencil** $\{W\}$. A point A of V is said to be a *base point* of the pencil $\{W\}$ *if every* $W_{\mathfrak{p}}$ *is on* A. Let A be a point of V and let $Q(A)=\mathfrak{J}$ be the quotient ring of A. We consider the intersection $R=\mathfrak{J}\cap P$. We prove the following theorem:

A is a base point of $\{W\}$ *if and only if the ring R is an algebraic extension of* K. *If A is not a base point of* $\{W\}$, *then there exists one and only one* $W_{\mathfrak{p}}$ *which passes through A.*

Proof. Let us first assume that R contains elements which are transcendental over K, whence *R is of degree of transcendency* 1 *over* K. Since V is normal, $\mathfrak{J}$ is integrally closed in Σ. Consequently

(a) *R is integrally closed in* P.

Since every element of P is algebraic over R, it follows by (a) that

(b) P *is the quotient field of R.*

Let ω be an arbitrary element of R. Since $\mathfrak{J}$ is the quotient ring of a zero-dimensional ideal, there exists a polynomial $f(\omega)$, with coefficients in K, such that $f(\omega)$ is a non-unit in $\mathfrak{J}$. Hence $1/f(\omega)\notin\mathfrak{J}$, whence $1/f(\omega)$ is an element of P which is not in R. Therefore

(c) *R is a proper subring of* P.

The non-units of $\mathfrak{J}$ form an ideal. Hence also

(d) *The non-units of R form an ideal.*

Since P is a field of algebraic functions of one variable, the properties (a), (b), (c), (d) of R imply that *R is a valuation ring.* Let $\mathfrak{p}$ be the corresponding prime divisor of P/K.

If *W is a member of the pencil and if* $W\neq W_{\mathfrak{p}}$, *then W does not pass through A*. For, let Γ be any irreducible component of W. The quotient ring $Q(\Gamma)$ is the valuation ring of the valuation defined by Γ (see Footnote 3). Since $W\neq W_{\mathfrak{p}}$, the intersection $Q(\Gamma)\cap P$ is different from R (this intersection is the valuation ring associated with a prime divisor of P/K different from $\mathfrak{p}$). Hence $Q(\Gamma)\cap P\supset R$, whence $Q(\Gamma)\supset\mathfrak{J}$. *Therefore A is not on* Γ. Since this holds true for any component Γ of W, A is not on W, as was asserted.

The variety $W_{\mathfrak{p}}$ *passes through A*. Let t be a non-unit of R. Then t is also a non-unit of $\mathfrak{J}$. Therefore at least one prime factor in the numerator of the principal divisor (t) must arise from a minimal prime of $\mathfrak{J}$. This minimal prime corresponds to an irreducible $(r-1)$-dimensional variety Γ passing

through A. The order of t along Γ is positive, whence the valuation of Σ/K defined by Γ induces in P/K a non-trivial valuation. The corresponding prime divisor of P/K must coincide with $\mathfrak{p}$, since $Q(\Gamma)\supset Q(A)\supset R$. Hence Γ is a component of $W_{\mathfrak{p}}$, and thus $W_{\mathfrak{p}}$ is on A.

Let us suppose now that R is an algebraic extension of K. Let $\mathfrak{p}$ be an *arbitrary* prime divisor of P/K. Let t be an element of P which has positive order at $\mathfrak{p}$ and nonpositive order at any other prime divisor of P/K. It is well known that such elements t exist. The element t is transcendental over K. Hence $1/t\notin R$, and consequently $1/t\notin\mathfrak{J}$. This implies that one of the prime factors in the numerator of the principal divisor (t) must arise from an irreducible subvariety Γ (of dimension $r-1$) which passes through A. Since t has positive order along Γ, it follows as before that Γ is a component of some member of the pencil, and this member must be $W_{\mathfrak{p}}$, since $\mathfrak{p}$ is the only prime divisor of P/K at which t has positive order. *Hence $W_{\mathfrak{p}}$ is on A.* Since $\mathfrak{p}$ was an arbitrary prime divisor of P/K, we conclude that A is a base point of the pencil $\{W\}$. This completes the proof of our theorem.

It is not difficult to see that *the base points of the pencil $\{W\}$ form an algebraic subvariety of V, of dimension $\leqq r-2$.* To see this, we fix an element t in P which is transcendental over K. Let $M_1,\cdots,M_\mu$ and $N_1,\cdots,N_\nu$ be the irreducible $(r-1)$-dimensional subvarieties of V along which t has respectively positive or negative order, and let H be the intersection of the two varieties $M_1+\cdots+M_\mu$ and $N_1+\cdots+N_\nu$. Clearly, H is of dimension $\leqq r-2$. We assert that *H is the base locus of the pencil $\{W\}$.* Namely, if A is a base point of $\{W\}$, then, by the preceding proof, neither t nor $1/t$ can belong to the quotient ring $Q(A)$, since both t and $1/t$ are transcendentals over K. Consequently the point A must lie on at least one of the varieties M_i and also on at least one of the varieties N_j. Therefore A is on H. Conversely, assume that A is on H, and let, say, A be on M_1 and N_1. It is clear that M_1 belongs to some member $W_{\mathfrak{p}}$ of the pencil, such that t has positive order at $\mathfrak{p}$. Similarly N_1 belongs to some member $W_{\mathfrak{p}'}$ of the pencil such that t has negative order at $\mathfrak{p}'$. Hence the point A lies on two *distinct* members of the pencil, $W_{\mathfrak{p}}$ and $W_{\mathfrak{p}'}$, and consequently A is a base point.

Remark. *Pencil with fixed components.* Let Γ_0 be a fixed $(r-1)$-dimensional subvariety of V and let $\{W\}$ be a pencil. It is convenient to regard also the collection of varieties $W'=W+\Gamma_0$ as a pencil. All members of this new pencil $\{W'\}$ have a fixed component, namely Γ_0. The base locus of $\{W'\}$ is naturally no longer of dimension $\leqq r-2$, since Γ_0 itself is a part of the base locus.

The pencils considered heretofore are pencils without fixed components.

II. The theorem of Bertini for pencils

9. Composite pencils. Let $\{W\}$, $\{W^*\}$ be two distinct pencils on V, both free from fixed components, and let P, P* be the corresponding subfields of Σ, of degree of transcendency 1 over K.

DEFINITION. *The pencil* $\{W\}$ *is composite with the pencil* $\{W^*\}$ *if* P *is a subfield of* P^*.

The geometric significance of this definition is straightforward. If P is a subfield of P^*, then P^* is an algebraic extension of P, since both have degree of transcendency 1 over K. Given a prime divisor $\mathfrak{p}$ of P/K, it factors in P^*/K:

$$\mathfrak{p} = \mathfrak{p}_1^{*\rho_1} \cdots \mathfrak{p}_s^{*\rho_s}, \tag{7}$$

where, if $\mathfrak{p}$ is of degree d and $\mathfrak{p}_i^*$ is of degree d_i^*, then $\rho_1 d_1^* + \cdots + \rho_s d_s^* = \nu d$, $\nu = [P^*: P]$. Consider the varieties $W_{\mathfrak{p}}$, $W_{\mathfrak{p}_i^*}^*$. If Γ^* is an irreducible component of $W_{\mathfrak{p}^*}$, then Γ^* is also a component of $W_{\mathfrak{p}}$, since $\mathfrak{p}_i^*$ lies over $\mathfrak{p}$. Moreover, if Γ^* occurs in $W_{\mathfrak{p}_i^*}^*$ to the multiplicity h_i, it must occur in $W_{\mathfrak{p}}$ to the multiplicity $h_i\rho_i$, since if an element of P has order 1 at $\mathfrak{p}$, it has order ρ_i at $\mathfrak{p}_i^*$. Finally, if Γ is an irreducible component of $W_{\mathfrak{p}}$ and if $\mathfrak{P}$ is the corresponding prime divisor of Σ, then $\mathfrak{P}$ induces in P the prime divisor $\mathfrak{p}$. Therefore $\mathfrak{P}$ must induce in P^* one of the prime divisors $\mathfrak{p}_i^*$, and consequently Γ is a component of one of the varieties $W_{\mathfrak{p}_i^*}^*$. *We conclude that*

$$W_{\mathfrak{p}} = \rho_1 W_{\mathfrak{p}_1^*}^* + \cdots + \rho_s W_{\mathfrak{p}_s^*}^*,$$

i.e., *if the pencil* $\{W\}$ *is composite with the pencil* $\{W^*\}$, *then every* W *consists of a certain number of members of the pencil* $\{W^*\}$.

In general this number $\rho_1 + \cdots + \rho_s$ depends on $W_{\mathfrak{p}}$. But if K is algebraically closed, then $d = d_i^* = 1$, *whence* $\rho_1 + \cdots + \rho_s = \nu$, the *relative degree* $[P^*: P]$[10]. For particular prime divisors $\mathfrak{p}$ it may happen that $\rho_1 + \cdots + \rho_s = 1$, i.e., that $W_{\mathfrak{p}}$ itself is a member of the pencil $\{W^*\}$. This can happen even for infinitely many prime divisors $\mathfrak{p}$. But it is easily seen that *there always exist infinitely many prime divisors* $\mathfrak{p}$ *for which* $\rho_1 + \cdots + \rho_s > 1$, so that *in a composite pencil there always exist infinitely many* $W_{\mathfrak{p}}$ *which are reducible over* K[11].

---

[10] There is one case in which our definition of composite pencils must be slightly modified. It is the case in which Σ is a field of functions of one variable. Since Σ itself is of degree of transcendency 1, it can be taken as field P and it therefore defines a pencil of zero-dimensional varieties on the curve V. This pencil is nothing else than the Riemann surface R of Σ, and every other pencil on V would be, according to our definition, composite with R. To avoid this undesirable conclusion, one must define composite pencils on an algebraic curve (i.e., composite *involutions of sets of points*) by the condition that the field P defining the pencil is not a maximal subfield of Σ.

[11] **Proof.** Let x be a transcendental element of P and let $\mathfrak{S}$ and $\mathfrak{S}^*$ denote the rings of integral functions of x in P and in P^* respectively. Let ω be a primitive element of $\mathfrak{S}^*$, with respect to P, which is not algebraic over K, and let D be the discriminant of the base $1, \omega, \cdots, \omega^{\nu-1}$. Let $\mathfrak{p}^*$ be a prime $\mathfrak{S}^*$-ideal at which the residue of ω is an element of K and such that $D \not\equiv 0$ $(\mathfrak{p}^*)$. There exist infinitely many such prime ideals $\mathfrak{p}^*$ since K contains infinitely many elements and since the numbers of prime divisors of P^* which are at infinity with respect to $\mathfrak{S}^*$ is finite. Let finally $\mathfrak{p}^* \cap \mathfrak{S} = \mathfrak{p}$. The assumption $D \not\equiv 0$ $(\mathfrak{p})$ implies that $1, \omega, \cdots, \omega^{\nu-1}$

The condition that $\{W\}$ is composite with $\{W^*\}$ implies, at any rate, that *the field* P *is not maximally algebraic in* Σ, since P^* is a proper algebraic extension of P. Conversely, if P is not maximally algebraic in Σ, and if P^* denotes the relative algebraic closure of P in Σ, then P^* is of degree of transcendency 1 over K, and the pencil $\{W\}$ is composite with pencil $\{W^*\}$ defined by P^*. Hence, a *pencil* $\{W\}$ is *composite* (with some other pencil) *if and only if the field* P *which defines it is not maximally algebraic in* Σ.

10. **Theorem of Bertini (for pencils).** *If a pencil* $\{W\}$, *free from fixed components, is not composite, then all but a finite number of members of the pencil are irreducible varieties* (*over* K).

For the proof we first establish two very simple lemmas, probably well known.

LEMMA 2. *If* P *is a maximally algebraic subfield of a field* Σ *and if* $\Sigma^* = \Sigma(x_1, \cdots, x_m)$ *is a pure transcendental extension of* Σ, *then* $P(x_1, \cdots, x_m)$ *is maximally algebraic in* Σ^*.

Proof. It is clearly sufficient to prove the lemma for $m=1$. Let then $\Sigma^* = \Sigma(x)$, x a transcendental over Σ, and let $t = f(x)/g(x)$ be an element of $\Sigma(x)$ which is algebraic over $P(x)$. We assume that $f(x)$ and $g(x)$ are relatively prime (in $\Sigma[x]$). We have to prove that t is an element of $P(x)$. Consider first the case $t \in \Sigma$ (whence f and g are elements of Σ). Since t is algebraic over $P(x)$, it must be already algebraic over P, since otherwise x would be algebraic over $P(t)$, and this is impossible, since $P(t) \subset \Sigma$. Hence $t \in P$, since P is maximally algebraic in Σ.

Let us now assume that t is not in Σ. Then t, being a transcendental over Σ, is *a fortiori* transcendental over P. But t is by hypothesis algebraic over $P(x)$. *Consequently* x *is algebraic over* $P(t)$. Let $f(x) = a_0x^n + \cdots + a_n$, $g(x) = b_0x^n + \cdots + b_n$; a_0, b_0 not both zero, and $a_i, b_j \in \Sigma$. Then

$$x^n(b_0t - a_0) + \cdots + (b_nt - a_n) = 0,$$

is the *irreducible* equation for x over $\Sigma(t)$ (since $f(x)$ and $g(x)$ are relatively prime in $\Sigma[x]$). The n roots of this equation are among the conjugates of x over $P(t)$, since $P(t) \subset \Sigma(t)$. *Consequently, the quotients*

$$\zeta_i = \frac{b_it - a_i}{b_0t - a_0}, \qquad i = 1, 2, \cdots, n,$$

are algebraic over $P(t)$.

Without loss of generality we may assume that $b_0 \neq 0$ (we may replace t by $1/t$). If we write down the algebraic equation for ζ_i over $P(t)$, then after

---

constitute an $\mathfrak{S}_\mathfrak{p}$-module basis for $\mathfrak{S}^*_{\mathfrak{p}^*}$. Hence if $\omega = c(\mathfrak{p})$ and if Δ, Δ^* denote the residue fields at $\mathfrak{p}$ and $\mathfrak{p}^*$ respectively ($\Delta = \mathfrak{S}/\mathfrak{p}$, $\Delta^* = \mathfrak{S}^*/\mathfrak{p}^*$), then $\Delta^* = \Delta(c)$, i.e., $\Delta^* = \Delta$, since $c \in K \subset \Delta$. This of course implies that $\mathfrak{S}^*\mathfrak{p} \neq \mathfrak{p}^*$, because $\mathfrak{S}^*\mathfrak{p} = \mathfrak{p}^*$ would imply the relation $[\Delta^* : \Delta] = \nu$.

clearing the denominators we get an algebraic equation for t over Σ. This must be an identity, since t is transcendental over Σ. If we equate to zero the leading coefficient of this equation in t, we find that b_i/b_0 *is algebraic over* P. But since P is maximally algebraic in Σ, it follows that b_i/b_0 *is an element of* P. Hence we may write $g(x)=b_0\cdot g_1(x)$, where $g_1(x)\in \mathrm{P}[x]$.

If also $a_0\neq 0$, then replacing t by $1/t$ we conclude in the same fashion that $f(x)=a_0 f_1(x), f_1(x)\in \mathrm{P}[x]$. We have now

$$t=\frac{a_0}{b_0}\cdot\frac{f_1(x)}{g_1(x)}\cdot$$

Since t is algebraic over $\mathrm{P}(x)$ and since $f_1(x)/g_1(x)\in \mathrm{P}(x)$, it follows that a_0/b_0 is algebraic over $\mathrm{P}(x)$. But since $a_0/b_0\in\Sigma$, it follows by the preceding case that $a_0/b_0\in \mathrm{P}$, whence $t\in \mathrm{P}(x)$. If $a_0=0$, then we replace t by $c+t$ where $c\in \mathrm{P}$ and $c\neq 0$. Then $c+t$ is a quotient of two polynomials, $(cg+f)/g$, *of the same degree*. By the preceding case $a_0\neq 0$, it follows that $c+t\in \mathrm{P}(x)$, whence also $t\in \mathrm{P}(x)$. This completes the proof of the lemma.

LEMMA 3. *Let* $f(x_1, \cdots, x_n)$ *be an irreducible polynomial in* $\mathrm{P}[x_1, \cdots, x_n]$, *where* P *is a field, and let* Σ *denote the field* $\mathrm{P}(\xi_1, \cdots, \xi_n)$ *defined by the equation* $f(\xi_1, \cdots, \xi_n)=0$. *If* P *is maximally algebraic in* Σ, *then the polynomial* $f(x_1, \cdots, x_n)$ *is absolutely irreducible.*

Proof. It is sufficient to prove that f remains irreducible for any finite extension Ω of P. The lemma is trivial when the field Ω is a pure transcendental extension of P: $\Omega=\mathrm{P}(u_1, \cdots, u_m)$. Moreover, in this case, if Σ^* denotes the field $\Omega(\xi_1, \cdots, \xi_n)$, then Σ^* is a pure transcendental extension of Σ, namely: $\Sigma^*=\Sigma(u_1, \cdots, u_m)$. Since P is maximally algebraic in Σ, it follows, by the preceding lemma, that Ω is maximally algebraic in Σ^*. We conclude that *it is sufficient to prove the lemma for finite algebraic extensions of* P. But in this case the lemma follows from a more general result concerning the behavior of prime ideals under algebraic extensions of the ground field (Zariski [5, Theorem 3, p. 198]).

11. **Proof of the theorem of Bertini.** The notation being that of §3, let ζ be any element of Σ^* and let

$$F(\{u_{ij}\}, \{\eta_i\}, \zeta) = 0 \tag{8}$$

be the irreducible equation for ζ over the field P^* $(=\mathrm{P}(\{u_{ij}\}, \{\eta_i\}))$. We may assume that F is an irreducible polynomial in $\mathrm{P}[\{u_{ij}\}, \{\eta_i\}, \zeta]$. Since Σ^* is a pure transcendental extension of Σ and since, by hypothesis, P is maximally algebraic in Σ, it follows, by Lemma 2, that P is also maximally algebraic in Σ^*. Hence, *a fortiori*, P is maximally algebraic in the field $\mathrm{P}(\{u_{ij}\}, \{\eta_i\}, \zeta)$, a subfield of Σ^*. By Lemma 3, we conclude that *the polynomial* F *is absolutely irreducible*:

Let $a, b, \cdots$ be the coefficients of the general polynomial in the arguments $\{u_{ij}\}, \{\eta_i\}, \zeta$, of the same degree as F ($a, b, \cdots$ are indeterminates). Let $a_0, b_0, \cdots$ be the corresponding coefficients of F ($a_0, b_0, \cdots$ are elements of P). There exists a finite number of finite sets of polynomials in $a, b, \cdots$, with rational coefficients, say

$$\{G_{i1}(a, b, \cdots), G_{i2}(a, b, \cdots), \cdots\}, \qquad i = 1, 2, \cdots, \rho,$$

with the following property, that if $F^*(\{u_{ij}\}, \{\eta_i\}, \zeta)$ is a polynomial, of the same degree as F, with coefficients $\alpha, \beta, \cdots$ in some field Ω, a necessary and sufficient condition that F^* be reducible in some extension of Ω is that $G_{i1}(\alpha, \beta, \cdots) = G_{i2}(\alpha, \beta, \cdots) = \cdots = 0$, for at least one value of i ($i = 1, 2, \cdots, \rho$). Since F is absolutely irreducible, it follows that for each i at least one of the elements

$$G_{i1}(a_0, b_0, \cdots), G_{i2}(a_0, b_0, \cdots), \cdots$$

is different from zero. These are elements of P. If then $\mathfrak{p}$ is a prime divisor of P/K, and if $a_0^*, b_0^*, \cdots$ denote the residues of $a_0, b_0, \cdots$ at $\mathfrak{p}$, then in the first place for all but a finite number of divisors $\mathfrak{p}$ the residues $a_0^*, b_0^*, \cdots$ will be all finite. In the second place, if an additional *finite* set of prime divisors $\mathfrak{p}$ is avoided, the expressions $G_{i1}(a_0^*, b_0^*, \cdots), G_{i2}(a_0^*, b_0^*, \cdots)$ will not be all zero, for any $i = 1, 2, \cdots, \rho$. Hence the resulting relation between the B^*-residues u_i^*, η_i^*, ζ^*:

$$F^*(\{u_{ij}^*\}, \{\eta_i^*\}, \zeta^*) = 0 \tag{9}$$

remains absolutely irreducible, and *in particular is irreducible over the residue field* Δ of $\mathfrak{p}$. Here B^* is the valuation of Σ^* defined by one of the divisors $\mathfrak{P}_i^*$ in (2), say by $\mathfrak{P}_1^*$.

Let now ζ be a primitive element of Σ^* over P^*. Then F is of degree ν in ζ, where $\nu = [\Sigma^* : P^*]$. Again, with the exception of a finite number of divisors $\mathfrak{p}$, the equation (9) will also be of degree ν in ζ^*. Since this equation is irreducible, it follows that the relative degree of the residue field Δ^* of $\mathfrak{P}_1^*$ with respect to the field $\Delta(\{u_i^*\}, \{\eta_i^*\})$ is at least ν. But this last field is the residue field of the prime divisor $\mathfrak{p}^*$ of P^*/K^*, and the relative degree in question cannot therefore exceed the relative degree $[\Sigma^* : P^*]$. Hence

$$[\Delta^* : \Delta(\{u_{ij}^*\}, \{\eta_i^*\})] = \nu. \tag{10}$$

Now the relative degree $[\Delta : K]$ is equal to d, where d is the degree of $\mathfrak{p}$. Since the u_{ij}^*, η_i^* are algebraically independent over Δ, it follows that $\Delta(\{u^*\}, \{\eta_i^*\})$ is also of relative degree d over the field $K(\{u_{ij}^*\}, \{\eta_i^*\})$. Hence, by (10),

$$[\Delta^* : K(\{u_{ij}^*\}, \{\eta_i^*\})] = \nu d.$$

But the left-hand side of this relation is the relative degree d_1 of $\mathfrak{P}_1^*$. Hence

$d_1=\nu d$. Comparing with the relation (6) of §7, we conclude that $m=1$, $h_1=1$, whence $W_{\mathfrak{p}}=\Gamma_1$, and $W_{\mathfrak{p}}$ is an irreducible subvariety of V. Since this holds true for all but a finite number of prime divisors of P/K, the proof of the theorem of Bertini is complete.

12. **Absolutely irreducible members of a non-composite pencil.** In this section we wish to prove the following theorem:

If K *is not maximally algebraic in* Σ, *then every irreducible subvariety of* V *is absolutely reducible. If* K *is maximally algebraic in* Σ, *and if* $\{W\}$ *is a non-composite pencil free from fixed components, then a* $W_{\mathfrak{p}}$ *is absolutely reducible if the degree of* $\mathfrak{p}$ *is* >1, *while, with a finite number of exceptions, all* $W_{\mathfrak{p}}$ *corresponding to prime divisors of degree* 1 *are absolutely irreducible.*

For the proof of this theorem, we first recall the definition of absolute irreducibility. An irreducible variety V in an $S_n(x_1, \cdots, x_n)$, over a ground field K, is given by a prime ideal $\mathfrak{p}$ in the polynomial ring $\mathfrak{J}=K[x_1, \cdots, x_n]$. The variety V is *absolutely irreducible* if $\mathfrak{p}$ remains prime under *any* extension K^* of the ground field K; i.e., if $\mathfrak{J}^*\mathfrak{p}$ is prime, $\mathfrak{J}^*=K^*[x_1, \cdots, x_n]$. In the contrary case, V is *absolutely reducible*. We now derive a necessary and sufficient condition for absolute irreducibility:

LEMMA 4. *A necessary and sufficient condition in order that* V *be absolutely irreducible is that* K *be maximally algebraic in the field* Σ *of rational functions on* V[12].

Proof. *The condition is sufficient.* If K^* is an algebraic extension of K, then the sufficiency of the condition follows from the quoted theorem in [5, p. 198]. If K^* is a pure transcendental extension of K, $K^*=K(u_1, \cdots, u_m)$, $\mathfrak{J}^*\mathfrak{p}$ is prime unconditionally. Moreover, if $\mathfrak{J}^*\mathfrak{p}=\mathfrak{p}^*$, then the quotient field Σ^* of $\mathfrak{J}^*/\mathfrak{p}^*$ coincides with the field $\Sigma(u_1, \cdots, u_m)$, a pure transcendental extension of Σ. By Lemma 2, if K is maximally algebraic in Σ, K^* will be maximally algebraic in Σ^*. Hence $\mathfrak{p}^*$ remains prime under any algebraic extension of K^*.

The condition is necessary. Let θ be an element of Σ which is algebraic over K, and let $F(\theta)=0$ be the irreducible equation for θ over K. Let $\theta_1, \cdots, \theta_{g-1}$ be conjugates of θ over K, where g is the degree of F. We take as field K^* the field $K(\theta, \theta_1, \cdots, \theta_{g-1})$. Since $\theta\in\Sigma$, there must exist a quotient ϕ/ψ $(\phi, \psi\in K[x_1, \cdots, x_n])$, $\psi\not\equiv 0$ $(\mathfrak{p})$, such that $\psi^g F(\phi/\psi)\equiv 0$ $(\mathfrak{p})$. This equation factors in $\mathfrak{J}^*$: $(\phi-\theta\psi)(\phi-\theta_1\psi)\cdots(\phi-\theta_{g-1}\psi)\equiv 0$ $(\mathfrak{J}^*\mathfrak{p})$. Now, by hypothesis, $\mathfrak{J}^*\mathfrak{p}$ is prime. Hence one of the factors $\phi-\theta_i\psi$ $(\theta_0=\theta)$ must be in $\mathfrak{J}^*\mathfrak{p}$. Since $\mathfrak{J}^*\mathfrak{p}$ is invariant under all automorphisms of K^* over K, it follows that all the g factors $\phi-\theta_i\psi$ must be in $\mathfrak{J}^*\mathfrak{p}$. This implies that g is 1, since $\theta_i\neq\theta_j$ if $i\neq j$ and since $\psi\not\equiv 0$ $(\mathfrak{p})$. Hence $\theta\in K$, and this shows that K is maximally algebraic in Σ.

(12) In Lemma 3 we have proved the sufficiency of this condition in the special case of principal ideals (f).

We now pass to the proof of the theorem stated at the beginning of this section.

Let us first assume that K is not maximally algebraic in Σ (whence V itself is absolutely reducible). *Since V is normal*, the relative algebraic closure $\overline{K}$ of K in Σ must be contained in the *integrally closed* ring $\mathfrak{o}=K[\xi_1, \cdots, \xi_n]$. Hence $\overline{K}$ is also contained in the quotient ring $\mathfrak{o}/\mathfrak{p}$ of any prime $\mathfrak{o}$-ideal $\mathfrak{p}$. Hence K is not maximally algebraic in the field of rational functions on the subvariety defined by $\mathfrak{p}$. Thus we conclude that in the present case *V does not carry at all absolutely irreducible subvarieties*. Note that the hypothesis that V is normal was essential in the proof.

We now consider the more interesting case in which the ground field K is maximally algebraic in Σ. Let $\mathfrak{p}$ be a prime divisor of P/K, of degree $d>1$. We show that *not only $W_\mathfrak{p}$ but also any irreducible component Γ of $W_\mathfrak{p}$ is absolutely reducible*. Namely, the field Δ of rational functions on Γ is the residue field of the prime divisor $\mathfrak{P}$ defined by Γ (since V is normal). Since $\mathfrak{P}$ induces in the field P the given prime divisor $\mathfrak{p}$, Δ contains the residue field of $\mathfrak{p}$. But since $d>1$, this last residue field is a proper algebraic extension of K. Hence K is not maximally algebraic in Δ, and Γ is absolutely reducible.

There remains to consider the prime divisors $\mathfrak{p}$ of degree 1. For these it is not difficult to see that the construction used in the proof of the theorem of Bertini actually leads to varieties $W_\mathfrak{p}$ which are not only irreducible over K but are also absolutely irreducible. In fact, in the first place we have the absolutely irreducible equation (9). The coefficients of this equation are elements of the residue field of $\mathfrak{p}$, i.e., in the present case, elements of K. The absolute irreducibility of (9) implies, by Lemma 4, that K *is maximally algebraic in the field* $K(\{u_{ij}^*\}, \{\eta_i^*\}, \zeta^*)$. In the second place, we may assume (avoiding a finite set of prime divisors $\mathfrak{p}$) that the equation (9) is of degree ν in ζ, where $\nu=[\Sigma^*:P^*]$. Then the field $K(\{u_{ij}^*\}, \{\eta_i^*\}, \zeta^*)$ coincides with the residue field Δ^* of $\mathfrak{P}_1^*$ (since this residue field is at most of relative degree ν with respect to the residue field of $\mathfrak{p}^*$, and since this last residue field is $K(\{u_{ij}^*\}, \{\eta_i^*\})$). We conclude that K is maximally algebraic in Δ^*. Since the field Δ of rational functions on the *irreducible* $W_\mathfrak{p}$ is a subfield of Δ^*, it follows that K is maximally algebraic in Δ and that consequently $W_\mathfrak{p}$ is absolutely irreducible (Lemma 4).

III. Linear systems of arbitrary dimension

13. **Definitions.** *A pencil $\{W\}$ is linear if the corresponding field* P *is a simple transcendental extension of* K (if $\{W\}$ has fixed components, P is the field which defines the pencil obtained from $\{W\}$ by deleting these fixed components).

Let us assume that $\{W\}$ is linear, and let $P=K(t)$, t a transcendental over K, $t\in\Sigma$. From our definition of a pencil, given in §2, it follows that our linear pencil $\{W\}$ can be obtained in the following fashion. Let $\mathfrak{p}$ be a prime

divisor of P/K and let us assume that $\mathfrak{p}$ is not the divisor at infinity of (t). Let $f(t^*)=0$ be the irreducible equation over K, of degree d, satisfied by the residue t^* of t at $\mathfrak{p}$. Here d is the degree of $\mathfrak{p}$. The principal divisor $(f(t))$ in Σ can be written in the form:

$$(f(t)) = \mathfrak{A}_{\mathfrak{p}}/\mathfrak{A}_{\infty}^{d}, \tag{11}$$

where $\mathfrak{A}_{\mathfrak{p}}$, $\mathfrak{A}_{\infty}$ are integral divisors (of the first kind) and $\mathfrak{A}_{\infty}$ is the denominator of the principal divisor (t) and is independent of $\mathfrak{p}$. *The member $W_{\mathfrak{p}}$ of the pencil $\{W\}$, outside of possible fixed components, is the subvariety of V which corresponds to the integral divisor $\mathfrak{A}_{\mathfrak{p}}$.* If $\mathfrak{p}$ is the prime divisor where t becomes infinite, then $\mathfrak{A}_{\mathfrak{p}}=\mathfrak{A}_{\infty}$ and $d=1$.

Suppose now that we consider only those prime divisors of P/K *which are of degree* $d=1$. Then $f(t)=c_0t+c_1$, $c_0, c_1\in$K, and we will have

$$(c_0t + c_1) = \frac{\mathfrak{A}_{\mathfrak{p}}}{\mathfrak{A}_{\infty}}.$$

This suggests a possible definition of a linear pencil which is different from ours. Namely one could define $\{W\}$ as the set of all $W_{\mathfrak{p}}$ for which $\mathfrak{p}$ is of degree 1. This is the customary definition in classical algebraic geometry. Unless K is algebraically closed, these usual pencils are proper subsets of our pencils. They are defined by the linear one-parameter family of functions: c_0t+c_1([13]). We use this customary procedure as a basis for the definition of linear systems. This definition is well known. We consider namely a finite set of functions in Σ, say $t_1, \cdots, t_s$, such that $1, t_1, \cdots, t_s$ are linearly independent over K. For arbitrary constants $c_0, c_1, \cdots, c_s$ in K, not all zero, let

$$c_0 + c_1t_1 + \cdots + c_st_s = \mathfrak{A}_{(c)}/\mathfrak{A}_{\infty},$$

where $\mathfrak{A}_{(c)}$ and $\mathfrak{A}_{\infty}$ are integral divisors (of the first kind) and where $\mathfrak{A}_{\infty}$ is independent of the c's. Let $W_{(c)}$ be the $(r-1)$-dimensional subvariety of V which corresponds to $\mathfrak{A}_{(c)}$. The totality of all $W_{(c)}$, as the constants c_i vary in K, is called *a linear system of dimensions s*, and is denoted by $|W|$. It is clear that $\mathfrak{A}_{\infty}$ itself is a particular $\mathfrak{A}_{(c)}$, namely for $c_0=1, c_1=\cdots=c_s=0$. It is also clear that we can always assume that no prime factor of $\mathfrak{A}_{\infty}$ occurs in *all* $\mathfrak{A}_{(c)}$ and that by this condition $\mathfrak{A}_{\infty}$ is uniquely determined. *The system $|W|$ is then free from fixed components.* If $\mathfrak{A}_{\infty}$ is chosen in this fashion, and if $\mathfrak{L}$ is an arbitrary *fixed* integral divisor, then $\mathfrak{A}_{(c)}\cdot\mathfrak{L}/\mathfrak{A}_{\infty}\cdot\mathfrak{L}$ is another representation of the principal divisor $(c_0+c_1t+\cdots+c_st)$, but it is not in reduced form. The system $|W|$ defined by this representation would have as fixed component the $(r-1)$-dimensional subvariety which corresponds to $\mathfrak{L}$.

([13]) All the results proved in the preceding sections continue to hold for these pencils, except the property that through each point of V there passes at least one $W_{\mathfrak{p}}$.

We note that if $|W|$ is free from fixed components and if

$$\mathfrak{A}_\infty = \mathfrak{P}_{\infty 1}^{h_1} \cdots \mathfrak{P}_{\infty m}^{h_m},$$

then each t_i has at $\mathfrak{P}_{\infty j}$ order $\geqq -h_j$, and at least one function t_i must have at $\mathfrak{P}_{\infty j}$ order exactly equal to $-h_j$.

One more remark. If we pass from $t_1, \cdots, t_s$ to another set of functions $\tau_1, \cdots, \tau_s$, by a nonsingular transformation of the form

$$\tau_i = \frac{d_{i0} + d_{i1}t_1 + \cdots + d_{is}t_s}{d_{00} + d_{01}t_1 + \cdots + d_{0s}t_s},$$

then the functions $1, \tau_1, \cdots, \tau_s$ define the same linear system $|W|$. Namely, we will have

$$(b_0 + b_1\tau_1 + \cdots + b_s\tau_s) = \frac{\mathfrak{A}_{(c')}}{\mathfrak{A}_{(d_{0j})}},$$

where $c_j' = b_0d_{0j} + b_1d_{1j} + \cdots + b_sd_{sj}$.

14. **Composite linear systems.** Let $\Omega = K(t_1, \cdots, t_s)$ *and let us assume that* Ω *is of degree of transcendency* 1 *over* K. Let us see what geometric property of $|W|$ corresponds to this assumption. We shall also assume that $|W|$ is free from fixed components.

Let $\{Z\}$ be the pencil (free from fixed components) defined by the field Ω. Let $t = c_0 + c_1t_1 + \cdots + c_st_s$, where $c_0, \cdots, c_s$ are arbitrary but fixed constants. The field $K(t) = P$ defines a pencil $\{W'\}$, free from fixed components. It is clear that each W' is either a W, or becomes a W after a certain fixed component M_0 is added to each W'. Namely, if t has order $-q_i$ at $\mathfrak{P}_{\infty i}$ $(i = 1, 2, \cdots, m)$, and if $\Gamma_{\infty i}$ is the irreducible V_{r-1} which corresponds to $\mathfrak{P}_{\infty i}$, then $W' + (h_1 - q_1)\Gamma_{\infty 1} + \cdots + (h_m - q_m)\Gamma_{\infty m}$ is a W, for each member W' of the pencil $\{W'\}$.

Let the pencil $\{W' + M_0\}$ be denoted by $\{W\}$. We have that P either coincides with Ω or is a subfield of Ω. *Let us first assume* that $\{W\}$ *has no fixed components.* We assert that *in this case* P *is a proper subfield of* Ω, *provided* $s > 1$.

To show this, let us assume the contrary. Each t_i is then an element of $K(t)$, say $t_i = f_i(t)/g_i(t)$, where f_i and g_i are assumed to be relatively prime. Let $t - d$ be a factor of $g_i(t)$, if g_i is of degree $\geqq 1$. Let $\mathfrak{P}$ be a prime divisor of Σ/K which occurs in the numerator of the principal divisor $(t - d)$. Since $\{W\}$ has no fixed components, t, and hence also $t - d$, becomes infinite along each of the varieties $\Gamma_{\infty 1}, \cdots, \Gamma_{\infty m}$. Consequently $\mathfrak{P} \neq \mathfrak{P}_{\infty i}$, $i = 1, 2, \cdots, m$. On the other hand, since $t - d$ is a factor of $g_i(t)$, t_i is infinite at the divisor $\mathfrak{P}$, and this is impossible since t_i can become infinite only at $\mathfrak{P}_{\infty 1}, \cdots, \mathfrak{P}_{\infty m}$.

Hence $g_i(t)$ *is of degree zero*, i.e., $t_i = f_i(t)$. Let ν_i be the degree of $f_i(t)$. Again, since $\{W\}$ is free from fixed components, the order of t at $\mathfrak{P}_{\infty j}$ is ex-

actly $-h_j$. Hence the order of t_i at $\mathfrak{P}_{\infty j}$ is exactly $-h_j\nu$. But this order is at most $-h_j$. Hence $\nu \leqq 1$. Thus each t_i is of the form: $t_i = d_{i0} + d_{i1}t$, and if $s > 1$ this is impossible, since $1, t_1, \cdots, t_s$ are linearly independent over K. This proves our assertion.

Since $\{W\}$ has no fixed components, every W' is a W, and since Ω is a proper algebraic extension of P, $\{W\}$ *is composite with the pencil* $\{Z\}$. We have therefore proved that each W decomposes into a certain number of varieties Z, provided W belongs to some subpencil of $|W|$ which is free from fixed components. But it is clear that there exists such a pencil for each W. Namely, let $W = W_{(c)}$ and let $W_{(c')}$ be a W which has no components in common with $W_{(c)}$. We put

$$t = \frac{c_0 + c_1 t_1 + \cdots + c_s t_s}{c_0' + c_1' t_1 + \cdots + c_s' t_s}.$$

It is then clear that the pencil determined by the field $K(t)$ is contained in the linear system $|W|$, is free from fixed components, and that $W_{(c)}$ and $W_{(c')}$ are members of that pencil.

In conclusion, we have proved that if $s > 1$ and if $K(t_1, \cdots, t_s)$ is of degree of transcendency 1 over K, then each member of the linear system $|W|$ (free from fixed components) decomposes into a certain number of varieties Z. We say that $|W|$ *is composite with the pencil* $\{Z\}$.

15. **The theorem of Bertini for linear systems.** We shall say that a linear system $|W|$ is *reducible* if *each* member W of the system is reducible *over* K. We shall say that $|W|$ is *absolutely reducible*, if each W is absolutely reducible.

THEOREM OF BERTINI. *A reducible linear system* $|W|$, *free from fixed components, is necessarily composite with a pencil. If* K *is maximally algebraic in* Σ, *then the assumption that* $|W|$ *is reducible can be replaced by the weaker assumption that* $|W|$ *is absolutely reducible.*

For linear systems of dimension 1 this theorem has already been proved in §§11 and 12. In order to prove this theorem for system of dimension $s > 1$, we first prove a lemma.

LEMMA 5. *Let* Σ *be a field of algebraic functions of* r *independent variables,* $r > 1$, *and let the ground field* K *be maximally algebraic in* Σ. *If* $x_1, \cdots, x_\rho$, $1 < \rho \leqq r$ *are algebraically independent elements of* Σ (*over* K) *and if we put*

$$\bar{x}_i = c_{i1}x_1 + \cdots + c_{i\rho}x_\rho, \qquad i = 1, 2, \cdots, \rho - 1,$$

then for non-special constants c_{ij} *in* K, *the field* $K(\bar{x}_1, \cdots, \bar{x}_{\rho-1})$ *is maximally algebraic in* Σ.

Proof. If the lemma is true for $\rho = 2$, then for non-special constants c_1, c_2 the field $\bar{K} = K(c_1x_1 + c_2x_2)$ is a maximally algebraic subfield of Σ, and the de-

gree of transcendency of Σ over $\overline{K}$ is $r-1$. If, say, $c_1 \neq 0$, then $x_2, \cdots, x_\rho$ are algebraically independent over $\overline{K}$, and hence replacing K by $\overline{K}$ we achieve a reduction from ρ to $\rho-1$. Therefore it is sufficient to prove the lemma for $\rho=2$.

Let Σ' be the relative algebraic closure of the field $K(x_1, x_2)$ in Σ. We may conduct the proof under the assumption that $\Sigma'=\Sigma$, for if $K(c_1x_1+c_2x_2)$ is maximally algebraic in Σ', then it is also maximally algebraic in Σ. Hence we may assume that $r=2$ and that therefore Σ is an algebraic extension of $K(x_1, x_2)$. Since we are dealing with fields of algebraic functions, Σ is a finite extension of $K(x_1, x_2)$.

Let $\bar{x}=x_1+cx_2$, $c \in K$, and let Ω_c denote the relative algebraic closure of the field $K(\bar{x})$ in Σ. We shall prove that *for all but a finite number of elements c in* K *the field* Ω_c *coincides with* $K(\bar{x})$. This result will establish the lemma.

Let $\Sigma_c = \Omega_c(x_2)$. We have

$$K(x_1, x_2) \subseteq \Sigma_c \subseteq \Sigma.$$

Since Σ is a *finite* algebraic extension of $K(x_1, x_2)$ and since for each c in K the field Σ_c is between $K(x_1, x_2)$ and Σ, it follows that there is only a *finite* number of *distinct* fields Σ_c[14]. Therefore, for all but a finite number of elements c in K, it is true that for a given c in K there exists another element d in K, $d \neq c$, such that $\Sigma_c = \Sigma_d$. We proceed to prove that *for any such element c in* K, *the field* Ω_c *coincides with* $K(\bar{x})$.

We may identify x_1+cx_2 and x_1+dx_2 with x_1 and x_2 respectively, since $c \neq d$. The fields Ω_c and Ω_d are now the relative algebraic closures in Σ of the fields $K(x_1)$ and $K(x_2)$ respectively. They shall be denoted by Ω_1 and Ω_2 respectively. Similarly we put $\Sigma_1 = \Omega_1(x_2)$, $\Sigma_2 = \Omega_2(x_1)$. We have, by hypothesis, $\Sigma_1 = \Sigma_2 = \Sigma^*$. The field K was assumed to be maximally algebraic in Σ. Hence K is also maximally algebraic in Ω_2. Since Σ^* $(=\Sigma_2)$ is a pure transcendental extension of Ω_2, it follows, by Lemma 2 of §10, that $K(x_1)$ *is maximally algebraic in* Σ^*. But $\Sigma^* = \Sigma_1 \supset \Omega_1$ and Ω_1 is algebraic over $K(x_1)$. Hence $\Omega_1 = K(x_1)$, as was asserted.

16. **Proof of the theorem of Bertini.** By a linear homogeneous transformation on $t_1, \cdots, t_s$ we may arrange matters so that t_1 becomes infinite on $\Gamma_{\omega i}$ $(i=1, 2, \cdots, m)$ to the highest order h_i. We consider the pencil $\{W\}$ contained in the system $|W|$ and defined by the functions $1, t_1+ct_i$, where $c \in K$. For all but a finite number of constants c in K, t_1+ct_i becomes infinite on $\Gamma_{\omega i}$ to the order h_i. Avoiding the exceptional constants c, we may therefore assert that the pencil $\{W\}$ is free from fixed components and consequently coincides with the pencil determined by the field $K(t_1+ct_i)$. However, the pencil $\{W\}$ contains only those varieties $W_\mathfrak{p}$ which correspond to prime divisors $\mathfrak{p}$ of

(14) Note that K is of characteristic zero and hence that Σ is a separable extension of $K(x_1, x_2)$. If we did not assume that K is of characteristic zero we could not assert that there is only a finite number of fields between $K(x_1, x_2)$ and Σ.

$K(t_1+ct_i)$ *which are of degree* 1. Let K^* be the relative algebraic closure of K in Σ, and let $\{W^*\}$ be the pencil defined by the field $K^*(t_1+ct_i)$. Again we include in $\{W^*\}$ only those varieties $W^*_{\mathfrak{p}^*}$ which correspond to prime divisors $\mathfrak{p}^*$ of $K^*(t_1+ct_i)$ which are of degree 1. If we identify the fields $K(t_1+ct_i)$, $K^*(t_1+ct_i)$ with the fields P, P^* respectively, of §9, we notice that if $\mathfrak{p}$ is of degree 1, then $\mathfrak{p}$ extends to a unique prime divisor of $\mathfrak{p}^*$ of P^*, i.e., we will have in (7): $s=1$, $\rho_1=1$. *Hence each member of the pencil* $\{W\}$ *is also a member of the pencil* $\{W^*\}$. By hypothesis, each W is reducible over K. Hence, by the theorem of Bertini for pencils, *the field* $K^*(t_1+ct_i)$ *is not maximally algebraic in* Σ. The same conclusion holds true if $K^*=K$ and if each W is absolutely reducible, in view of the theorem of §12. Since this is true for all but a finite number of constants c in K, it follows, by Lemma 5, that t_1 and t_i are algebraically dependent over K^*, and hence also over K. Since this holds for $i=2, \cdots, s$ and since t_1 is a transcendental over K, it follows that the field $K(t_1, \cdots, t_s)$ is an algebraic extension of $K(t_1)$ and is therefore of degree of transcendency 1, q.e.d.

References

1. W. Krull, *Idealtheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 4, no. 3.

2. B. L. van der Waerden, *Moderne Algebra*, vol. 2.

3. B. L. van der Waerden, *Über lineare Scharen von reduziblen Mannigfaltigkeiten*, Mathematische Annalen, vol. 113 (1936).

4. O. Zariski, *Some results in the arithmetic theory of algebraic varieties*, American Journal of Mathematics, vol. 61 (1939).

5. O. Zariski, *Algebraic varieties over ground fields of characteristic zero*, American Journal of Mathematics, vol. 62 (1940).

The Johns Hopkins University,
Baltimore, Md.
California Institute of Technology,
Pasadena, Calif.

SOLUTION OF THE INVERSE PROBLEM OF THE CALCULUS OF VARIATIONS

BY
JESSE DOUGLAS

PART I. INTRODUCTION

1. Formulation and background. The problem indicated in the title is one of the most important hitherto unsolved problems of the calculus of variations, namely:

Given any family of ∞^{2n} curves (paths) in $(n+1)$-dimensional space (x, y_j), $(j=1, \cdots, n)$, as represented by a system of differential equations

$$y_i'' = F_i(x, y_j, y_j') \qquad (i = 1, \cdots, n); \tag{1.1}$$

to determine whether these curves can be identified with the totality of extremals of some variation problem

$$\int \phi(x, y_j, y_j')dx = \min, \tag{1.2}$$

and in the affirmative case to find all the corresponding functions ϕ.

The present paper solves this problem for the most important and interesting case of 3-dimensional space $(n=2)$, where the given family consists of ∞^4 curves defined by differential equations of the form

$$y'' = F(x, y, z, y', z'), \qquad z'' = G(x, y, z, y', z'), \tag{1.3}$$

and the variation problem sought for is of the form

$$\int \phi(x, y, z, y', z')dx = \min. \tag{1.4}$$

Our essential results and methods have already been published in two preliminary notes[1].

Basically, our procedure consists in an application of the Riquier theory of systems of partial differential equations to a certain linear differential system $\mathfrak{S}$ on which the inverse problem can be made to depend. This differential system has already appeared—derived in a different way—in the interesting work, of little more than a decade ago, by D. R. Davis on the inverse problem[2]; but, as he stated, its general solution—even existence-theoretically—presented difficulties which he could not overcome.

Presented to the Society, January 1, 1941; received by the editors March 13, 1940.

(1) Numbers 8 and 9 of the list of references at the end of §2.

(2) See numbers 6 and 7 of the list of references at the end of §2.

In this paper we succeed in obtaining a complete solution of the differential system $\mathfrak{S}$ in the following sense. We determine in all cases whether or not a solution exists corresponding to given (F, G)—that is, a classification is made of all curve families (F, G) into "extremal" and "nonextremal." In each case of consistency of the system $\mathfrak{S}$ we derive an equivalent *completely integrable* differential system $[\mathfrak{S}]$ which determines ϕ. This system $[\mathfrak{S}]$ is either given explicitly—that is, the coefficients in its equations (all linear) are expressed as functions of the partial derivatives of (F, G)—or else (when this seems to lead to calculations too lengthy to be worth while) we describe precisely those differentiations and algebraic processes (additions and multiplications), finite in number, which will furnish this completely integrable system. The explicit formulation of $[\mathfrak{S}]$ is found to be feasible in all the important cases. Standard criteria of the theory of differential systems then enable us to determine the degree of generality of the solution ϕ, that is, the number and nature of the arbitrary functions or constants which are involved.

It may be emphasized that the coordinates (x, y_i) are by no means necessarily cartesian, but may be perfectly general point coordinates. Indeed, as is evident *a priori*, the fundamental group of our problem is that of *all* (analytic) point transformations (or, otherwise interpreted, all coordinate transformations); for if the curves (1.1) are the extremals of the integral (1.2), this relation obviously continues to subsist, after arbitrary transformation of the variables (x, y_i), between the transformed curves and the transformed integral.

We assume the functions F_i, ϕ and any others that may enter into our investigations to be *analytic* about a particular linear element $(x^{(0)}, y_j^{(0)}, y_j'^{(0)})$ which we may take to be located at the origin along the x-axis, that is, with all coordinates zero.

The principal formulas of a tensor form of treatment of the inverse problem, based on parametric representation, have been given in one of our preliminary notes(3). In the present exposition the non-parametric form of representation, as in all the preceding formulas, will be employed throughout.

Analytic formulation. Analytically expressed, our inverse problem consists in the solution for ϕ, as unknown function, of the Euler-Lagrange equations

$$\frac{\partial \phi}{\partial y_i} - \frac{d}{dx}\frac{\partial \phi}{\partial y_i'} = 0 \qquad (i = 1, \cdots, n). \tag{1.5}$$

Here the operator d/dx denotes total differentiation with respect to x along an arbitrary curve of the given family; that is, as applied to any function of x, y_j, y_j',

$$\frac{d}{dx} \equiv \frac{\partial}{\partial x} + y_j'\frac{\partial}{\partial y_j} + F_j\frac{\partial}{\partial y_j'}, \tag{1.6}$$

(3) Number 8 of the list at the end of §2.

where, as throughout this paper, *the summation convention applies to repeated indices.*

In expanded form, the Euler-Lagrange equations are, then:

$$\frac{\partial \phi}{\partial y_i} - \left(\frac{\partial}{\partial x} + y_j' \frac{\partial}{\partial y_j} + F_j \frac{\partial}{\partial y_j'}\right)\frac{\partial \phi}{\partial y_i'} = 0 \qquad (i = 1, \cdots, n), \tag{1.7}$$

where F_j are given and ϕ is to be found. Further, these equations must be solvable uniquely as a linear algebraic system for the quantities F_j or y_j'', in order that the differential equations of the given curve family may have the prescribed form (1.1). Therefore a solution of (1.7) is required such that

$$\text{Det } |\phi_{ij}| \neq 0, \tag{1.8}$$

where

$$\phi_{ij} \equiv \frac{\partial^2 \phi}{\partial y_i' \partial y_j'}. \tag{1.9}$$

The determinant (1.8) is, of course, the Hessian of ϕ.

For $n = 1$, or a 2-dimensional containing space, the solution of the inverse problem is classic, being given in the standard textbooks(4). The system (1.7) then consists of but a single equation, whose solvability for ϕ is assured by known existence theorems. Thus, *any* given family of ∞^2 curves in the plane, as defined by a differential equation

$$y'' = F(x, y, y'), \tag{1.10}$$

can be regarded as the totality of extremals of a problem

$$\int \phi(x, y, y')dx = \text{min.}$$

Indeed, ϕ can be explicitly determined by quadratures if the differential equation (1.10) has been integrated: $y = f(x, a, b)$. The formula for ϕ involves essentially one arbitrary function of two arguments(5), so that the class of integrals corresponding to a given curve family assigned as extremals is quite extensive.

The difficulty of the problem resides in the case $n \geqq 2$, that is, for a 3- or higher dimensional space. Then the number of equations (1.7) to be satisfied exceeds the number (one) of unknown functions, and for this reason it has always been presumed (correctly) that for *arbitrarily* given F_j the equations (1.7) have no solution except the trivial one of any total derivative:

(4) See Bolza, *Lectures on the Calculus of Variations*, New York, 1931, pp. 31–32. Cf. also the reference to the work of Darboux in §2.

(5) In this count we omit the arbitrary function $\nu(x, y)$ in the exact differential $d\nu(x, y)$ that may be added to $\phi(x, y, y')dx$.

$$\phi = \frac{d}{dx} v(x, y_j) \equiv \frac{\partial v}{\partial x} + \frac{\partial v}{\partial y_j} y_j',$$

which evidently does not satisfy the condition (1.8) since each element (1.9) is equal to zero. In other words, the presumption is that a given family (1.1) of ∞^{2n} curves in $(n+1)$-space is, in general, not an extremal family.

However, it may be an indication of the difficulty of discussing the system (1.7) directly that hitherto no *proof* has ever been given of this surmisal. That is to say, no example has ever been given of a $2n$-parameter curve family in $(n+1)$-space which is certainly not an extremal family(6).

In the course of the general solution of the inverse problem that occupies this paper, many examples of nonextremal families are found and proved to be such; for instance, in xyz-space

$$y'' = y^2 + z^2, \qquad z'' = y \tag{1.11}$$

is a non-extremal family(7), so is

$$y'' = y^2 + z^2, \qquad z'' = 0(8); \tag{1.12}$$

and as many others as are desired can be constructed.

2. **Historical survey.** Darboux in his geometrical treatise(9) stated and solved the inverse problem for the 2-dimensional case.

G. Hamel, in 1903, gave a solution of a particular case of the inverse problem in 3-space where the assigned extremals are the totality of straight lines.

Along another line, Jacobi, in his fundamental memoir of 1837 on the calculus of variations, proved that if $\delta E = 0$ is the variational equation for the Euler equation (of $2n$th order):

$$E \equiv \frac{\partial \phi}{\partial y} - \frac{d}{dx}\frac{\partial \phi}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial \phi}{\partial y''} - \cdots + (-1)^n \frac{d^n}{dx^n}\frac{\partial \phi}{\partial y^{(n)}} = 0, \tag{2.1}$$

corresponding to a 2-dimensional problem of any order n

$$\int \phi(x, y, y', y'', \cdots, y^{(n)})dx = \min, \tag{2.2}$$

then δE is always *self-adjoint* as a linear differential expression in δy.

In 1897, A. Hirsch proved that, conversely, if

$$\omega(x, y, y', \cdots, y^{(2n)})$$

is any differential expression whose variation $\delta\omega$ (a linear differential expres-

(6) Except in the preliminary notes to this paper, cited in Footnote 1.
(7) Cf. §18.
(8) Cf. §17.
(9) See, throughout, the list of references at the end of this section.

sion in δy) is self-adjoint, then ω must be identifiable with the Euler expression E corresponding to some integrand ϕ, as defined by (2.1). Later, J. Kürschak extended this result to expressions involving partial derivatives.

D. R. Davis, in a Chicago dissertation of about a decade ago, proved the corresponding converse theorem for a system of n differential expressions $\omega_i(x, y_j, y_j', y_j'')$: if their variations $\delta\omega_i$ form a self-adjoint linear differential system in δy_j, then $\phi(x, y_j, y_j')$ exists such that the ω_i are the first members of Euler's equations for $\int\phi dx = \min$. On this basis, Davis transformed the inverse problem in $(n+1)$-space into the determination of multipliers $P_{ij}(x, y_k, y_k')$ for the given differential equations (1.1) such that the following differential expressions:

$$E_i = P_{ij}\{F_j(x, y_k, y_k') - y_j''\},$$

have variations δE_i which are self-adjoint.

By imposing this condition of self-adjointness, Davis obtained for the multipliers P_{ij} precisely the differential system $\mathfrak{S}$ which we shall derive in an entirely different way in the present paper (Part II)([10]). In attempting, however, to solve the system $\mathfrak{S}$ even existence-theoretically, Davis met with difficulties which, as he stated, he was unable to overcome. He therefore contented him self with the study of certain three particular examples([11]).

The present paper *accomplishes the solution of the system* $\mathfrak{S}$ *from the existence-theoretic standpoint.* This includes the determination in all the various cases of the existence and generality (in terms of arbitrary functions and constants) of the solution. In the case of existence of a solution, a *completely integrable* differential system is given, either explicitly or implicitly, for the determination of this solution.

Incidentally, we derive the fundamental differential system $\mathfrak{S}$ from an entirely new point of view, based on certain identities which we have found to be obeyed by the Euler expressions. Self-adjointness thus plays no rôle whatever in our theory and may be omitted entirely from consideration.

References

1. G. Darboux, *Leçons sur la Théorie Générale des Surfaces*, Paris, 1894, §§604, 605.

2. G. Hamel, *Über die Geometrieen in denen die Geraden die Kürzesten sind*, Mathematische Annalen, vol. 57 (1903), pp. 231–264.

([10]) It should be remarked that Davis does not explicitly include in the system $\mathfrak{S}$ the important inequation $LN - M^2 \neq 0$ (see formula (7.5)), which plays a very prominent role in our treatment of the system.

([11]) These are: (i) the straight lines, (ii) all circles orthogonal to a fixed plane, (iii) catenaries (cf. the example at the end of §12 of the present paper, where this case is treated by a different method). Davis does not seem to have noticed explicitly that (ii) is equivalent to (i) by a point transformation, so that these examples are essentially identical from the standpoint of our problem (cf. the fifth paragraph in §1).

3. C. G. J. Jacobi, *Zur Theorie der Variationsrechnung und der Differentialgleichungen*, Werke, vol. 4.

4. A. Hirsch, *Über eine charakteristische Eigenschaft der Differentialgleichungen der Variationsrechnung*, Mathematische Annalen, vol. 49 (1897), pp. 49–72.

5. J. Kürschak, *Über eine charakteristische Eigenschaft der Differentialgleichungen der Variationsrechnung*, Mathematische Annalen, vol. 60 (1905), pp. 157–165.

6. D. R. Davis, *The inverse problem of the calculus of variations in higher space*, these Transactions, vol. 30 (1928), pp. 710–736.

7. D. R. Davis, *The inverse problem of the calculus of variations in a space of* $(n+1)$ *dimensions*, Bulletin of the American Mathematical Society, vol. 35 (1929), pp. 371–380.

8. J. Douglas, *Solution of the inverse problem of the calculus of variations*, Proceedings of the National Academy of Sciences, vol. 25 (1939), pp. 631–637.

9. J. Douglas, *Theorems in the inverse problem of the calculus of variations*, ibid., vol. 26 (1940), pp. 215–221.

The last two are preliminary notes to the present paper.

3. **Statement of results.** Because of the number and variety of our results, it seems desirable to present them first before proceeding to their detailed development and proof. In this section we therefore collect the theorems whose establishment is the goal of the main body, Part IV, of this paper. We shall also give an example of each of the important types which arise in our classification.

Of central importance is a certain matrix

$$\Delta \equiv \begin{Vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{Vmatrix} \tag{3.1}$$

whose elements are known functions depending on the given curve family (F, G), being expressions in the partial derivatives of F, G. Explicitly:

$$\begin{aligned} A &= \frac{d}{dx}F_{x'} - 2F_x - \tfrac{1}{2}F_{x'}(F_{y'} + G_{x'}), \\ B &= -\frac{d}{dx}F_{y'} + \frac{d}{dx}G_{x'} + 2(F_y - G_x) + \tfrac{1}{2}(F_{y'} - G_{x'})(F_{y'} + G_{x'}), \\ C &= -\frac{d}{dx}G_{y'} + 2G_y + \tfrac{1}{2}G_{y'}(F_{y'} + G_{x'}), \end{aligned} \tag{3.2}$$

where for the 3-dimensional case given in (1.3), the differential operator d/dx of (1.6) is

$$\frac{d}{dx} \equiv \frac{\partial}{\partial x} + y'\frac{\partial}{\partial y} + z'\frac{\partial}{\partial z} + F\frac{\partial}{\partial y'} + G\frac{\partial}{\partial z'}. \tag{3.3}$$

A_1, B_1, C_1 are derived from A, B, C by the formulas

$$A_1 = \frac{dA}{dx} - F_{y'}A - \tfrac{1}{2}F_{z'}B,$$

$$B_1 = \frac{dB}{dx} - G_{y'}A - \tfrac{1}{2}(F_{y'} + G_{z'})B - F_{z'}C, \tag{3.4}$$

$$C_1 = \frac{dC}{dx} - \tfrac{1}{2}G_{y'}B - G_{z'}C,$$

while A_2, B_2, C_2 are derived from A_1, B_1, C_1 by the same formulas (recursion).

The *rank* of Δ is the principal basis of classification of curve families (F, G) with respect to our inverse problem; accordingly, we begin by a separation into cases according to this rank, which will be followed by the appropriate sub-classifications in the statement of our theorems.

Case I. $\|A\ B\ C\| = 0$, that is, $A = 0$, $B = 0$, $C = 0$.

Case II.

$$\left\| \begin{matrix} A & B & C \\ A_1 & B_1 & C_1 \end{matrix} \right\| = 0, \qquad \|A\ B\ C\| \neq 0.$$

Case III.

$$\begin{vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} = 0, \qquad \left\| \begin{matrix} A & B & C \\ A_1 & B_1 & C_1 \end{matrix} \right\| \neq 0.$$

Case IV.

$$\begin{vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} \neq 0.$$

Here, in writing a matrix $= 0$, we mean that each determinant resulting therefrom by the suppression of columns (only) is equal to zero, and $\neq 0$ means that at least one such determinant is not equal to zero.

It is seen by reference to the recursion formulas (3.4) that the cases thus described are precisely those of rank 0, 1 2, 3 of the matrix Δ, respectively.

Case II is the most difficult and varied in its results. Fundamental in its treatment is the quadratic equation

$$A\xi^2 + B\xi + C = 0, \tag{3.5}$$

whose roots we denote by λ, μ. It is important to distinguish the case of unequal roots from that of equal roots:

Case IIa. $B^2 - 4AC \neq 0$, $\lambda \neq \mu$;

Case IIb. $B^2 - 4AC = 0$, $\lambda = \mu$.

λ, μ are known functions of x, y, z, y', z' when the curve family is given, being algebraic functions of the partial derivatives of F, G.

The investigation of Case IIa leads to an interesting division into three subcases according to the following criteria:

Case IIa1. $\lambda\lambda_{z'}-\lambda_{y'}=0$, $\mu\mu_{z'}-\mu_{y'}=0$;

Case IIa2. $\lambda\lambda_{z'}-\lambda_{y'}=0$, $\mu\mu_{z'}-\mu_{y'}\neq 0$;

Case IIa3. $\lambda\lambda_{z'}-\lambda_{y'}\neq 0$, $\mu\mu_{z'}-\mu_{y'}\neq 0$.

For reasons that will appear in our later detailed discussion (cf. end §10), we shall refer to these cases respectively as *"separated," "semi-separated," and "non-separated."*

Case IIb will be subdivided into

Case IIb1. $\lambda\lambda_{z'}-\lambda_{y'}=0$;

Case IIb2. $\lambda\lambda_{z'}-\lambda_{y'}\neq 0$.

Case IIb1 will be further divided into

Case IIb1′. $\lambda\lambda_{z'}-\lambda_{y'}=0$, $(\mathrm{IX}')=0$;

Case IIb1″. $\lambda\lambda_{z'}-\lambda_{y'}=0$, $(\mathrm{IX}')\neq 0$:

where (IX') is a certain expression in the partial derivatives of F, G, given explicitly by formula (14.10) modified according to the remarks following formula (15.2).

In Case III, let Δ_1, Δ_2, Δ_3 denote the second order determinants contained in the matrix of A, B, C and A_1, B_1, C_1:

$$\Delta_1 = BC_1 - B_1C, \qquad \Delta_2 = CA_1 - C_1A, \qquad \Delta_3 = AB_1 - A_1B.$$

The vanishing or nonvanishing of

$$D = \Delta_1\Delta_3 - \Delta_2^2$$

is important in the discussion of Case III; accordingly, we subdivide this case:

Case IIIa. $D\neq 0$,

Case IIIb. $D=0$.

We shall indicate the generality of the solution of a given differential system by the notation

$$\infty^{m_1f(n_1)+m_2f(n_2)+\cdots+m_kf(n_k)+p}\ (12) \tag{3.6}$$

to represent the presence in the general solution of m_1 arbitrary functions of n_1 arguments, m_2 arbitrary functions of n_2 arguments, $\cdots$, m_k arbitrary functions of n_k arguments, and p arbitrary constants.

We remark that the extremals of $\int\phi dx=\min$ and $\int[\phi dx+d\nu(x,y,z)]=\min$, where ν denotes an arbitrary function, are evidently the same. For

$$\int_{P_1}^{P_2}[\phi dx+d\nu(x,y,z)]=\int_{P_1}^{P_2}\phi dx+\nu(x_2,y_2,z_2)-\nu(x_1,y_1,z_1),$$

(12) This symbol is due to E. Kasner, appearing in his review of Riquier's treatise on differential systems, Bulletin of the American Mathematical Society, vol. 19 (1913), p. 14.

where $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ are any two points of space. Since the difference between the values of the two integrals is independent of the path followed from P_1 to P_2, a minimizing path for one is also minimizing for the other.

It is also evident that ϕ and $c\phi$, where c denotes an arbitrary nonzero constant factor, have the same extremals.

In estimating the generality of the integrand $\phi(x, y, x, y', z')$ corresponding to given F, G, *we shall omit the arbitrary function* $\nu(x, y, z)$ *and the arbitrary constant* c, considering the class of integrands

$$c\phi(x, y, z, y', z') + \frac{d}{dx}\nu(x, y, z) \tag{3.7}$$

as defining essentially a single variation problem equivalent to $\int \phi dx = \min$.

We are now ready to state our results and to cite our illustrative examples. As has been said, in all cases where the given curve family is of extremal type, that is, a corresponding variation problem exists, we shall give (in Part IV) a completely integrable differential system for the determination of the integrand ϕ. All these systems are linear in the partial derivatives of the unknown functions.

THEOREM I. *In Case* I *the given curve family is always of extremal type, and the generality of the corresponding variation problem is expressed by the symbol* $\infty^{2f(3)+2f(2)}$.

This theorem is illustrated by the straight lines: $y'' = 0$, $z'' = 0$; also more generally by $y'' = f(z')$, $z'' = 0$, the function f being arbitrary.

THEOREM II. *In Case* IIa1 (*separated*) *the given curve family is of extremal type and the generality of the corresponding variation problem is* $\infty^{2f(2)}$. *The explicit determination of the integrand* ϕ *depends on the solution of two separate complete systems of linear partial differential equations of the first order*[13], *with respective unknown functions* $\rho(x, y, z, y', z')$, $\sigma(x, y, z, y', z')$; *there are three equations in each system.*

Example. $y'' = F(x, y, y')$, $z'' = G(x, z, z')$, with $B \neq 0$. The inverse problem here separates into one in the xy-plane and one in the xz-plane, whose solutions by addition give the integrand ϕ for the space problem.

THEOREM III. *In Case* IIa2 (*semi-separated*) *the solution of the inverse problem depends first on the solution for* σ *of a certain system of linear partial differential equations of first order in that unknown function, which system may be rendered complete in the usual way by the adjunction of alternants. Its solution is then substituted in a system of three linear partial differential equations of first*

[13] See the reference in Footnote 24.

order involving ρ together with σ, whereupon this system becomes a complete one for ρ. According to the number of linearly independent equations in the complete system for σ (4, 5, or 6), the generality of the integrand ϕ is $\infty^{1/(2)+1/(1)}$, or $\infty^{1/(2)}$, or else ϕ is nonexistent, that is, the given curve family is nonextremal.

Example. The ∞^4 catenaries which lie in planes perpendicular to the xz-plane (pictured as horizontal) and the directrix of each of which coincides with the trace of its plane upon the xz-plane. This family of catenaries is represented by the system of differential equations

$$y'' = \frac{1 + y'^2 + z'^2}{y}, \qquad z'' = 0.$$

The generality of the corresponding variation problem is $\infty^{1/(2)+1/(1)}$.

The discussion of the non-separated case, IIa3, is, in the nature of the matter, not as simple or neat as that of the others. Nevertheless, every eventuality is followed out in the detailed considerations given in §13. Even in the most favorable hypothetical possibility, the generality of the corresponding variation problem is certainly not higher than $\infty^{2/(1)}$, as will be seen by a review of §13.

THEOREM IV. *In Case* IIb1′ *the given curve family is of extremal type and the generality of the corresponding variation problem is $\infty^{2/(2)}$. The determination of ϕ depends on the solution of a complete system of three partial differential equations of first order for $\sigma(x, y, z, y', z')$ followed by the solution of a similar system for $\rho(x, y, z, y', z')$.*

Example. $y'' = z$, $z'' = 0$ (see §15).

THEOREM V. *In Case* IIb1″ *no corresponding variation problem exists.*

An example is given at the end of §15.

Case IIb2 is discussed in §16.

THEOREM VI. *In Case* IIIa *the solution of the inverse problem depends on a certain differential*

$$E_1dx + E_2dy + E_3dz + E_4dy' + E_5dz',$$

whose coefficients E_i are definite functions of the partial derivatives of F, G—indeed, rational functions having D for denominator. In case this differential is exact, the corresponding variation problem exists and is essentially unique(14); *in the contrary case, no corresponding variation problem exists.*

Example. $y'' = z^2$, $z'' = y^2$. The corresponding integral is

$$\int (y'z' + \tfrac{1}{3}y^3 + \tfrac{1}{3}z^3)dx,$$

(14) Cf. formula (3.7) and the associated remarks.

uniquely determined up to the possibility of the addition of an arbitrary exact differential and multiplication by an arbitrary constant factor (cf. formula 3.7).

THEOREM VII. *In Case* IIIb *no corresponding variation problem exists.*

Example. $y''=y^2+z^2$, $z''=0$.

THEOREM VIII. *In Case* IV *no corresponding variation problem exists.*

Example. $y''=y^2+z^2$, $z''=y$.

PART II. THE FUNDAMENTAL DIFFERENTIAL SYSTEM $\mathfrak{S}$

4. **Derivation of $\mathfrak{S}$.** The first step in our procedure is to make the solution of the original Euler-Lagrange equations (1.7) for the unknown function ϕ depend on an equivalent differential system $\mathfrak{S}$, more easily manageable, where the unknowns are the functions ϕ_{ij} defined by (1.9). We find it convenient to conduct the discussion for a general value of the dimensionality $n+1$, afterwards putting $n=2$.

Our derivation of the system $\mathfrak{S}$ is based on certain *new identities* which we establish involving the *Euler expressions*

$$\omega_i \equiv \frac{\partial \phi}{\partial y_i} - \frac{d}{dx}\frac{\partial \phi}{\partial y_i'} \qquad (i = 1, \cdots, n). \tag{4.1}$$

These identities are

$$\frac{\partial \omega_i}{\partial y_j'} + \frac{\partial \omega_j}{\partial y_i'} + 2\frac{d}{dx}\phi_{ij} + \frac{\partial F_k}{\partial y_j'}\phi_{ik} + \frac{\partial F_k}{\partial y_i'}\phi_{jk} = 0, \tag{4.2}$$

$$\begin{aligned}\frac{d}{dx}\left(\frac{\partial \omega_i}{\partial y_j'} - \frac{\partial \omega_j}{\partial y_i'}\right) - 2\left(\frac{\partial \omega_i}{\partial y_j} - \frac{\partial \omega_j}{\partial y_i}\right) + \frac{1}{2}\frac{\partial F_k}{\partial y_i'}\left(\frac{\partial \omega_j}{\partial y_k'} + \frac{\partial \omega_k}{\partial y_j'}\right) \\ - \frac{1}{2}\frac{\partial F_k}{\partial y_j'}\left(\frac{\partial \omega_i}{\partial y_k'} + \frac{\partial \omega_k}{\partial y_i'}\right) + A_{jk}\phi_{ik} - A_{ik}\phi_{jk} = 0,\end{aligned} \tag{4.3}$$

where

$$A_{jk} \equiv \frac{d}{dx}\frac{\partial F_k}{\partial y_j'} - 2\frac{\partial F_k}{\partial y_j} - \frac{1}{2}\frac{\partial F_m}{\partial y_j'}\frac{\partial F_k}{\partial y_m'}. \tag{4.4}$$

Their verification may be left to the reader with the remark that here and frequently afterwards in this paper it is helpful to observe the following "alternating relations" between differential operators:

$$\frac{\partial}{\partial y_j}\frac{d}{dx} - \frac{d}{dx}\frac{\partial}{\partial y_j} = \frac{\partial F_k}{\partial y_j}\frac{\partial}{\partial y_k'}, \tag{4.5}$$

$$\frac{\partial}{\partial y_j'}\frac{d}{dx} - \frac{d}{dx}\frac{\partial}{\partial y_j'} = \frac{\partial}{\partial y_j} + \frac{\partial F_k}{\partial y_j'}\frac{\partial}{\partial y_k'}. \tag{4.6}$$

If the Euler-Lagrange equations $\omega_i=0$ are satisfied, then the preceding identities imply the following equations obeyed by the functions ϕ_{ij}:

$$\frac{d}{dx}\phi_{ij}+\frac{1}{2}\frac{\partial F_k}{\partial y_j'}\phi_{ik}+\frac{1}{2}\frac{\partial F_k}{\partial y_i'}\phi_{jk}=0, \tag{4.7}$$

$$A_{jk}\phi_{ik}-A_{ik}\phi_{jk}=0. \tag{4.8}$$

We also have evidently from (1.9):

$$\frac{\partial \phi_{ij}}{\partial y_k'}=\frac{\partial \phi_{ik}}{\partial y_j'}, \tag{4.9}$$

and we take explicit notice of the symmetry of ϕ_{ij} in its indices:

$$\phi_{ij}=\phi_{ji}. \tag{4.10}$$

The differential system in ϕ_{ij} consisting of the equations (4.7)–(4.10) together with the inequation (1.8):

$$\text{Det}\,|\phi_{ij}|\neq 0, \tag{4.11}$$

is our fundamental differential system $\mathfrak{S}$.

It is to be observed that all the equations of $\mathfrak{S}$ are *linear*.

5. **Equivalence of the system $\mathfrak{S}$ to the Euler-Lagrange equations.** Conversely, let ϕ_{ij} be any system of functions of x, y_k, y_k' which obeys the system $\mathfrak{S}$.

Then (4.9) implies the existence of functions

$$\phi_i=\int \phi_{ij}dy_j', \tag{5.1}$$

since the differential under the integral sign is exact. We have $\partial\phi_i/\partial y_j'=\phi_{ij}$; therefore by (4.10), $\partial\phi_i/\partial y_j'=\partial\phi_j/\partial y_i'$, from which it is possible to conclude the exactness of the differential $\phi_i dy_i'$ and the existence of a function

$$\phi=\int \phi_i dy_i'. \tag{5.2}$$

Evidently the functions ϕ_{ij} are the second partial derivatives of ϕ, as required by (1.9).

The existence of a function ϕ whose second partial derivatives $\partial^2\phi/\partial y_i'\partial y_j'$ are precisely the functions ϕ_{ij} being thus assured, it is next evident that ϕ is determined only up to the addition of an arbitrary linear function of the y_k':

$$\phi=\bar{\phi}+\lambda_0+\lambda_k y_k' \tag{5.3}$$

where $\bar{\phi}$ denotes any fixed determination and λ_0, λ_k are functions of x, y_j.

Let functions $\bar{\omega}_i$ be deduced from any fixed determination $\bar{\phi}$ by formula (4.1). Then by the identities (4.2, 3) and the equations (4.7, 8), it follows that

$$\frac{\partial \bar{\omega}_i}{\partial y_j'} + \frac{\partial \bar{\omega}_j}{\partial y_i'} = 0, \tag{5.4}$$

$$\frac{d}{dx}\left(\frac{\partial \bar{\omega}_i}{\partial y_j'} - \frac{\partial \bar{\omega}_j}{\partial y_i'}\right) - 2\left(\frac{\partial \bar{\omega}_i}{\partial y_j} - \frac{\partial \bar{\omega}_j}{\partial y_i}\right) = 0. \tag{5.5}$$

Applying $\partial/\partial y_k'$ to (5.4), we get

$$\frac{\partial^2 \bar{\omega}_i}{\partial y_j' \partial y_k'} + \frac{\partial^2 \bar{\omega}_j}{\partial y_i' \partial y_k'} = 0;$$

and cyclic permutation of i, j, k gives three equations of this type, wherefrom

$$\frac{\partial^2 \bar{\omega}_i}{\partial y_j' \partial y_k'} = 0.$$

This expresses that $\bar{\omega}_i$ is linear in the variables y_k':

$$\bar{\omega}_i = \alpha_{ik} y_k' + \beta_i, \tag{5.6}$$

where α_{ik}, β_i are functions of x, y_j. Formula (5.4) then implies the skew-symmetry of α_{ij}:

$$\alpha_{ij} + \alpha_{ji} = 0. \tag{5.7}$$

Substituting (5.6) in (5.5), we find

$$\frac{\partial \alpha_{ij}}{\partial y_k} + \frac{\partial \alpha_{ki}}{\partial y_j} + \frac{\partial \alpha_{jk}}{\partial y_i} = 0, \tag{5.8}$$

$$\frac{\partial \alpha_{ij}}{\partial x} + \frac{\partial \beta_j}{\partial y_i} - \frac{\partial \beta_i}{\partial y_j} = 0. \tag{5.9}$$

If we extend the range $1, \cdots, n$ of the indices i, j in α_{ij} to include the index 0 by defining $\alpha_{i0} = \beta_i$, $\alpha_{0i} = -\beta_i$, $\alpha_{00} = 0$ (which agrees with (5.7)), then (5.9) becomes

$$\frac{\partial \alpha_{ij}}{\partial x} + \frac{\partial \alpha_{j0}}{\partial y_i} + \frac{\partial \alpha_{0i}}{\partial y_j} = 0; \tag{5.10}$$

we may say therefore that (5.8) holds for the range $0, 1, \cdots, n$ of the indices (provided, further, that y_0 is identified with x).

Let now $\bar{\phi}$ be replaced by the general determination ϕ according to (5.3). Then, as is readily calculated from (4.1), $\bar{\omega}_i$ is transformed into

$$\omega_i = \bar{\omega}_i - \left(\frac{\partial \lambda_i}{\partial x} - \frac{\partial \lambda_0}{\partial y_i}\right) - \left(\frac{\partial \lambda_i}{\partial y_k} - \frac{\partial \lambda_k}{\partial y_i}\right) y_k'. \tag{5.11}$$

By the substitution (5.6), we see that $\omega_i = 0$ if and only if

$$(5.12)\qquad \begin{aligned} &\frac{\partial \lambda_i}{\partial y_k} - \frac{\partial \lambda_k}{\partial y_i} = \alpha_{ik}, && (i, k = 1, \cdots, n), \\ &\frac{\partial \lambda_i}{\partial x} - \frac{\partial \lambda_0}{\partial y_i} = \beta_i = \alpha_{i0}, && (i = 1, \cdots, n). \end{aligned}$$

These conditions express that the *curl* of the unknown vector $\lambda = (\lambda_0, \lambda_1, \cdots, \lambda_n)$ shall be the known skew-symmetric tensor $\alpha_{\rho\sigma}$, $(\rho, \sigma = 0, 1, \cdots, n)$. It is well-known that the established relations (5.8), (5.10) are exactly necessary and sufficient in order that such a vector λ exist. λ is then determined up to the addition of an arbitrary *gradient* $(\partial v/\partial x, \partial v/\partial y_1, \cdots, \partial v/\partial y_n)$ where v is any function of x, y_j. That is, according to (5.3), *ϕ is determined up to the addition of an arbitrary total derivative as to x:*

$$\frac{d}{dx} v(x, y_k) = \frac{\partial v}{\partial x} + \frac{\partial v}{\partial y_k} y_k'.$$

As we have already remarked, it is evident *a priori* that we could not expect to determine ϕ any more closely than this, since the addition of an exact differential $dv(x, y_k)$ to ϕdx has no effect on the extremals of $\int \phi dx$.

In summary, we may state the following

PROPOSITION[15]. *To every solution ϕ_{ij} of the system $\mathfrak{S}$ there corresponds a solution ϕ of the Euler-Lagrange equations having ϕ_{ij} as the system of its second partial derivatives as to the $y''s$: $\partial^2\phi/\partial y_i' \partial y_j' = \phi_{ij}$. For a given solution ϕ_{ij}, the function ϕ can be found by successive quadratures[16], and is uniquely determined except for the inevitable addition of an arbitrary total derivative $dv(x, y_k)/dx$.*

In other words, the inverse problem of the calculus of variations is precisely equivalent to the solution for ϕ_{ij} of the differential system $\mathfrak{S}$.

PART III (§6). OUTLINE OF THE GENERAL THEORY OF DIFFERENTIAL SYSTEMS

The essence of this paper consists in an application of the general theory of differential systems, as presented in the standard treatises[17], to the particular system $\mathfrak{S}$. Referring the reader to these works, we shall presuppose the vocabulary, ideas, and facts of this theory. However, for purposes of ex-

[15] This proposition, as well as the proof here given (essentially), is contained in the papers of Davis referred to in §2.

[16] The determination of a vector, given its curl, can be effected by a series of quadratures.

[17] These are, principally: C. Riquier, *Les Systèmes d'Équations aux Dérivées Partielles*, Paris, 1910; M. Janet, *Leçons sur les Systèmes d'Équations aux Dérivées Partielles*, Paris, 1929; J. F. Ritt, *Differential Equations from the Algebraic Standpoint*, American Mathematical Society

position—particularly, to render the sequel intelligible to those not well acquainted with the theory of differential systems, as developed principally by Riquier—it seems desirable to give here a brief outline of the procedure for testing the consistency of any given differential system and, in case a solution exists, for obtaining an equivalent completely integrable system which will give this solution. The application of the general theory to our particular case is very considerably simplified by the circumstance that $\mathfrak{S}$ is linear.

By a *differential system* we understand any finite number of equations in which appear any finite number n of independent variables x_j, and any finite number m of unknown functions u_i of these variables, together with a certain finite number of partial derivatives of the functions u_i with respect to the x_j up to any order. Each member of each equation is supposed to be an analytic function of the variables x_j, u_i and those partial derivatives of the u_i which it involves effectively, this function being regular, that is, representable as a power-series, about certain initial values of all these variables. The system may also contain a certain finite number of *inequations*; that is, certain analytic functions of x_j, u_i and some of the partial derivatives of the u_i may be required to be $\neq 0$.

If a differential system is in a form solved for certain derivatives of the unknown functions u_i, we term any derivative of a function u_i which can be obtained by differentiation of a first member to any order with respect to the independent variables x_j a *principal derivative.* (This definition is interpreted to include the first members themselves among the principal derivatives, as their own derivatives of zeroth order.) All other partial derivatives of the functions u_i are termed *parametric.*

Riquier introduced the device of *cotes.* With each independent variable x_j is associated a composite integer with any fixed number s of components: $c_j \equiv (c_{j1}, c_{j2}, \cdots, c_{js})$, where each c_{jk} is an integer $\geqq 0$; this composite integer is what is called the "cote" of x_j. Similarly a cote is assigned to each unknown function u_i: $c_i' \equiv (c_{i1}', c_{i2}', \cdots, c_{is}')$. The cote of any partial derivative

$$\frac{\partial^{r_1+r_2+\cdots+r_n} u_i}{\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}}$$

is then, by definition,

$$\Gamma \equiv (\gamma_1, \gamma_2, \cdots, \gamma_s)$$

where $\gamma_k = c_{ik}' + r_1 c_{1k} + r_2 c_{2k} + \cdots + r_n c_{nk}$.

Colloquium Publications, vol. 14, New York, 1932; J. M. Thomas, *Differential Systems*, American Mathematical Society Publications, vol. 21, New York, 1937.

See also an exposition of the Riquier theory by Thomas, Annals of Mathematics, (2), vol. 30 (1929), p. 285 et seq.

Particularly relevant from our viewpoint are: Riquier, chap. 7; Janet, pp. 74–75; Ritt, chap. 9; Thomas, §§36, 37, 75, 76.

All the derivatives of all the unknown functions([18]) are arranged according to their cotes in "dictionary order," that is, given any two derivatives δu_i, $\delta' u_j$ with respective cotes

$$\Gamma \equiv (\gamma_1, \gamma_2, \cdots, \gamma_s), \qquad \Gamma' \equiv (\gamma_1', \gamma_2', \cdots, \gamma_s'),$$

δu_i is called "*higher*" than $\delta' u_j$ if the first of the differences

$$\gamma_1 - \gamma_1', \gamma_2 - \gamma_2', \cdots, \gamma_s - \gamma_s'$$

which is not zero is positive; if the first of these nonzero differences is negative, δu_i is called "*lower*" than $\delta' u_j$, or $\delta' u_j$ higher than δu_i. In case $\Gamma = \Gamma'$, that is, $\gamma_k = \gamma_k'$ for $k = 1, \cdots, s$, no distinction of higher and lower is created between δu_i and $\delta' u_j$. The cotes can always be assigned so that this eventuality never occurs, and the ordering of derivatives produced by the cotes is then said to be *complete*.

It is an essential condition, always supposed verified in the Riquier theory, that *the first cote c_{j1} of each independent variable x_j* be equal to 1.

Riquier terms a differential system $\mathfrak{S}$ *orthonomic* if it obeys the following three conditions:

1° it is solved for *distinct* derivatives of the unknown functions;

2° no second member contains any principal derivative effectively;

3° cotes have been assigned, in accordance with the stipulation $c_{j1} = 1$, so that each first member is *higher* than any derivative which appears effectively in the corresponding second member.

The *prolonged system* $\mathfrak{S}'$ of any given differential system $\mathfrak{S}$ is the infinite system which is derived from $\mathfrak{S}$ by subjecting each of its equations to every possible differentiation with respect to the independent variables x_j: $\partial^{r_1+r_2+\cdots+r_n}/\partial x_1^{r_1}\partial x_2^{r_2}\cdots\partial x_n^{r_n}$. In this differentiation each member of each equation of $\mathfrak{S}$ is considered as a composite function of the x_j through the intermediary of those partial derivatives of the functions u_i which appear in this member, and the differentiation is effected according to the composite function rule.

The important idea of a *passive* system $\mathfrak{S}$ is defined as follows. $\mathfrak{S}$ is called "passive" if its prolonged system $\mathfrak{S}'$ is equivalent to a system $\mathfrak{S}''$ which consists of a unique expression for each principal derivative in terms of a finite number of parametric derivatives. This means that the equations of $\mathfrak{S}'$ will be satisfied identically in the parametric derivatives if we substitute for each principal derivative its expression from $\mathfrak{S}''$.

A fundamental theorem of the Riquier theory is that *every differential system which is passive and orthonomic is completely integrable.* The meaning of this is that a unique solution of the given differential system exists which

([18]) The unknown functions will be regarded throughout as their own derivatives of zeroth order.

corresponds to initial data of the following nature: the values of all the parametric derivatives are assigned arbitrarily for any fixed initial values $x_j^{(0)}$ of the independent variables x_j, provided only that we have *convergence* of that part of the Taylor series of each unknown function u_i about $x_j^{(0)}$ which involves these parametric derivatives as coefficients. In other words, the initial values of all principal derivatives may be regarded as having been put equal to zero to give this "*initial determination*" of the functions u_i.

The same initial determination may be represented in the manner now to be described. It is a fact, proved in the general theory, that all parametric derivatives may be obtained from a certain finite number among them, called a *complete set*, as follows. With respect to each parametric derivative of the complete set, we distinguish the independent variables x_j according to a certain rule as *multipliers* and *non-multipliers*. Then each parametric derivative is obtainable from a unique one of the complete set by a unique differentiation with respect to its multipliers. From this, one can infer that the initial assignment of all parametric derivatives subject to the stated convergence condition is equivalent to assigning to each partial derivative δu_i of the complete set an arbitrary convergent Taylor series in the multipliers x_j of δu_i (power-series $P(x_j - x_j^{(0)})$ in the quantities $x_j - x_j^{(0)}$) and requiring δu_i to reduce to this series when the non-multipliers x_k are set equal to their initial values: $\delta u_i = P(x_j - x_j^{(0)})$ for each $x_k = x_k^{(0)}$. This type of initial condition then fixes a solution of the given differential system uniquely.

One may consult J. M. Thomas, loc. cit., §§75, 76, or Ritt, loc. cit., p. 139 for the notion of complete set, both of principal and of parametric derivatives, and for the definition of multipliers and non-multipliers of each derivative belonging to a complete set. It is convenient here, as throughout the theory of differential systems, to employ an obvious isomorphism between partial derivatives of a given function $u(x_1, x_2, \cdots, x_n)$ and algebraic monomials in the variables x_j: thus $\partial^{r_1+r_2+\cdots+r_n} u / \partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}$ corresponds to the monomial $x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$, and any successive operations of differentiation with respect to the independent variables x_j correspond to multiplication of the associated monomials.

The importance of the passive nature of a differential system makes it essential to have a criterion for the passivity of a given orthonomic differential system $\mathfrak{S}$. Each equation of the prolonged system $\mathfrak{S}'$ expresses a certain principal derivative in terms of parametric derivatives and of principal derivatives of lower cote; this follows from the orthonomic nature of $\mathfrak{S}$, which can be proved to extend to the prolonged system $\mathfrak{S}'$. The expressions for the principal derivatives thus furnished by $\mathfrak{S}'$ are said to be *direct*. It is much to be emphasized that a given principal derivative may occur as first member in many different equations of $\mathfrak{S}'$, that is, be obtainable in many different ways by differentiation of first members of $\mathfrak{S}$. For example, if $\partial^3 u / \partial x \partial y^2$, $\partial^3 u / \partial x^2 \partial y$, $\partial u / \partial x$ are first members of $\mathfrak{S}$, the derivative $\partial^5 u / \partial x^3 \partial y^2$ may

be obtained from these respectively by the differentiations $\partial^2/\partial x^2$, $\partial^2/\partial x \partial y$, $\partial^4/\partial x^2 \partial y^2$. Accordingly, any principal derivative may, and generally will, have various direct expressions.

Imagine the principal derivatives arranged in order of cote from lower to higher as first, second, third, $\cdots$: $P_1, P_2, P_3, \cdots$. Then the lowest principal derivative P_1 must, as can easily be shown, be a first member of $\mathfrak{S}$, and has therefore, by condition 1° for an orthonomic system, a unique expression in terms of parametric derivatives exclusively (the independent variables x_j are supposed to figure throughout as parameters). Consider next the second lowest principal derivative P_2; each of its direct expressions involves, besides parametric derivatives, only the lowest principal derivative P_1; if the expression just mentioned for this is substituted, then a certain number of expressions are obtained for P_2 in terms of parametric derivatives exclusively—these are termed the *ultimate expressions* of P_2. Next, the third lowest principal derivative P_3 has a certain number of direct expressions in terms of parametric derivatives and the two lower principal derivatives P_1, P_2. If in each of these direct expressions for P_3 we substitute the unique ultimate expression for P_1 and, in succession, all the different ultimate expressions already obtained for P_2, we get a certain number of ultimate expressions for P_3 in terms of parametric derivatives exclusively. Continuing step-by-step in this way, we obtain for each principal derivative P_k a certain number of ultimate expressions in terms of parametric derivatives.

A criterion for the passivity of the given differential system is that *all the various ultimate expressions of each principal derivative P_k shall be identical.*

This criterion may be considerably simplified by basing the test for passivity on only a certain finite number of principal derivatives, called *cardinal derivatives*, instead of the infinite set of all of them. If $\partial^{r_1+r_2+\cdots+r_n} u_i/\partial x_1^{r_1} \partial x_2^{r_2} \cdots \partial x_n^{r_n}$ and $\partial^{s_1+s_2+\cdots+s_n} u_i/\partial x_1^{s_1} \partial x_2^{s_2} \cdots \partial x_n^{s_n}$ are any two first members of $\mathfrak{S}$ which are partial derivatives of the same unknown function u_i, then the corresponding cardinal derivative is, by definition, $\partial^{t_1+t_2+\cdots+t_n} u_i/\partial x_1^{t_1} \partial x_2^{t_2} \cdots \partial x_n^{t_n}$ where each t_j is the greater of r_j and s_j([19]). Then, as proved by Riquier, the following is a necessary and sufficient condition for the passivity of $\mathfrak{S}$: *the various ultimate expressions for each cardinal derivative shall be identical.*

This is the form of test which will be used in the applications that follow.

If $\mathfrak{S}$ is not passive, the application of the preceding criterion, that is, the equating of every two different ultimate expressions for a cardinal derivative, will lead to relations among the parametric derivatives which are not verified identically. By solving these for the parametric derivatives of highest cote([20]), we obtain an orthonomic differential system $\mathfrak{S}_1$ equivalent to $\mathfrak{S}$ in the sense that any solution of the one must be a solution of the other.

([19]) In terms of the isomorphism between partial derivatives and monomials mentioned earlier in this section, cardinal derivative corresponds to least common multiple.

([20]) Cf. Janet, loc. cit., p. 75.

The passage from $\mathfrak{S}$ to $\mathfrak{S}_1$ clearly enlarges the class of principal derivatives and narrows that of parametric derivatives.

If we next apply our passivity test to $\mathfrak{S}_1$, then either it is verified that $\mathfrak{S}_1$ is passive, or else new equations not identically satisfied are obtained among parametric derivatives of $\mathfrak{S}_1$. In the latter case a system $\mathfrak{S}_2$ is derived as before, equivalent to $\mathfrak{S}$ and $\mathfrak{S}_1$, and with a still wider class of principal derivatives and a narrower one of parametric derivatives.

It is a standard fact of the theory of differential systems that this process, accompanied by successive enlargement of the class of principal derivatives and restriction of the class of parametric derivatives, *can go on for only a finite number k of steps*. That this is so is a direct consequence of the following theorem due to Riquier: *if in a sequence of monomials* $x_1'^{r_1}x_2'^{r_2}\cdots x_n'^{r_n}$ (*each* r_i *an integer* $\geqq 0$) *no monomial is a multiple of a predecessor, the sequence can contain only a finite number of terms*. (Cf. Janet, loc. cit., p. 11; Ritt, loc. cit., p. 135; and refer also to the isomorphism between partial derivatives and monomials mentioned earlier in this section.)

At the end of the finite process just described, either the conditions for passivity are found to be verified, or else we obtain a non-identical relation involving only the independent variables, or a relation contradicting one of the specified inequations of $\mathfrak{S}$. In the latter two eventualities the system $\mathfrak{S}$ is inconsistent with itself, or has no solutions; in the former case we obtain a passive orthonomic system $[\mathfrak{S}]=\mathfrak{S}_k$ equivalent to $\mathfrak{S}$, which has therefore a unique solution corresponding to initial data of the type previously described; that is, $[\mathfrak{S}]$ is completely integrable.

The main body of this paper, which follows, is concerned with applying the general theory just outlined to the particular differential system $\mathfrak{S}$ on which depends the inverse problem of the calculus of variations.

Part IV. Solution of the differential system $\mathfrak{S}$

7. **The system $\mathfrak{S}$ in the three-dimensional case.** We have derived the differential system $\mathfrak{S}$ in §§4, 5 for a general value of the dimensionality $n+1$. For the three-dimensional inverse problem, which alone will be considered in this paper, $n=2$, and we adopt the notation $(x, y_1, y_2, y_1', y_2') = (x, y, z, y', z')$, $(F_1, F_2) = (F, G)$, so that, as already written in §1, the differential equations of the given curve family are

$$y'' = F(x, y, z, y', z'), \qquad z'' = G(x, y, z, y', z'), \tag{7.1}$$

while the variation problem sought for is

$$\int \phi(x, y, z, y', z')dx = \min. \tag{7.2}$$

For the functions ϕ_{ij} $(i, j = 1, 2)$ we write

$$L = \phi_{y'y'}, \qquad M = \phi_{y'z'}, \qquad N = \phi_{z'z'}. \tag{7.3}$$

The differential operator (1.6) is now

$$\frac{d}{dx} \equiv \frac{\partial}{\partial x} + y'\frac{\partial}{\partial y} + z'\frac{\partial}{\partial z} + F\frac{\partial}{\partial y'} + G\frac{\partial}{\partial z'}. \tag{7.4}$$

By comparison with (4.7)–(4.11), we see that the fundamental differential system $\mathfrak{S}$ now consists of the following six linear equations and quadratic inequation in L, M, N as unknown functions:

$$\begin{gathered}
\frac{dL}{dx} + F_{y'}L + G_{y'}M = 0, \\
\frac{dM}{dx} + \tfrac{1}{2}F_{z'}L + \tfrac{1}{2}(F_{y'} + G_{z'})M + \tfrac{1}{2}G_{y'}N = 0, \\
\frac{dN}{dx} + F_{z'}M + G_{z'}N = 0, \\
AL + BM + CN = 0, \\
L_{z'} = M_{y'}, \\
N_{y'} = M_{z'}, \\
LN - M^2 \neq 0.
\end{gathered} \tag{7.5}$$

In the purely algebraic equation (7.5_4) of the system, the coefficients A, B, C are, by (4.4),

$$\begin{aligned}
A &= \frac{d}{dx}F_{z'} - 2F_z - \tfrac{1}{2}F_{z'}(F_{y'} + G_{z'}), \\
B &= -\frac{d}{dx}F_{y'} + \frac{d}{dx}G_{z'} + 2(F_y - G_z) + \tfrac{1}{2}(F_{y'}^2 - G_{z'}^2), \\
C &= -\frac{d}{dx}G_{y'} + 2G_y + \tfrac{1}{2}G_{y'}(F_{y'} + G_{z'}).
\end{aligned} \tag{7.6}$$

Applying successively to (7.5_4) the operator d/dx, and using the first three equations of (7.5), we obtain the following two additional purely algebraic linear equations in L, M, N:

$$A_1L + B_1M + C_1N = 0, \tag{7.7}$$

$$A_2L + B_2M + C_2N = 0. \tag{7.8}$$

Here A_1, B_1, C_1 are derived from A, B, C by the formulas

$$A_1 = \frac{dA}{dx} - F_{y'}A - \tfrac{1}{2}F_{z'}B,$$

$$B_1 = \frac{dB}{dx} - G_{y'}A - \tfrac{1}{2}(F_{y'} + G_{z'})B - F_{z'}C, \tag{7.9}$$

$$C_1 = \frac{dC}{dx} - \tfrac{1}{2}G_{y'}B - G_{z'}C,$$

while A_2, B_2, C_2 are derived from A_1, B_1, C_1 by the same formulas (recursion).

Since $(L, M, N) \not\equiv (0, 0, 0)$—for the contrary would be inconsistent with the last condition of (7.5)—we infer from the system of linear equations (7.5_4), (7.7), (7.8) *the following property of all extremal curve families*:

$$\begin{vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} = 0. \tag{7.10}$$

This necessary condition gives a very easy way of constructing nonextremal curve families, namely, by the nonvanishing of the determinant just written. The readily verified example given in (1.11) was found in this manner[21].

8. **Case I. When the fundamental matrix is of rank zero.** As stated in §3, the principal basis of our classification into cases is the rank of the matrix

$$\Delta \equiv \left\| \begin{matrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{matrix} \right\|. \tag{8.1}$$

The first case to consider is that of rank zero, that is, when

$$A = 0, \quad B = 0, \quad C = 0. \tag{8.2}$$

It is evident by the recursion formulas (7.9) that every element of Δ is then zero; that is, the rank of Δ is indeed zero.

The fourth equation (7.5) now disappears identically, so that the system $\mathfrak{S}$ is now constituted by the remaining five equations and the inequation of (7.5). We proceed to consider $\mathfrak{S}$ from the standpoint of the Riquier theory.

This is the place to make the following preliminary remark. In the theory of the system $\mathfrak{S}$ we may use d/dx instead of $\partial/\partial x$, that is, we may take as fundamental the differential operators d/dx, $\partial/\partial y$, $\partial/\partial z$, $\partial/\partial y'$, $\partial/\partial z'$ instead of $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$, $\partial/\partial y'$, $\partial/\partial z'$. In doing this, it is important to have always in mind the following "alternating relations":

[21] Cf. §18.

$$
\begin{aligned}
&\frac{\partial}{\partial y}\frac{d}{dx}-\frac{d}{dx}\frac{\partial}{\partial y}=F_y\frac{\partial}{\partial y'}+G_y\frac{\partial}{\partial z'},\\
&\frac{\partial}{\partial z}\frac{d}{dx}-\frac{d}{dx}\frac{\partial}{\partial z}=F_z\frac{\partial}{\partial y'}+G_z\frac{\partial}{\partial z'},\\
&\frac{\partial}{\partial y'}\frac{d}{dx}-\frac{d}{dx}\frac{\partial}{\partial y'}=\frac{\partial}{\partial y}+F_{y'}\frac{\partial}{\partial y'}+G_{y'}\frac{\partial}{\partial z'},\\
&\frac{\partial}{\partial z'}\frac{d}{dx}-\frac{d}{dx}\frac{\partial}{\partial z'}=\frac{\partial}{\partial z}+F_{z'}\frac{\partial}{\partial y'}+G_{z'}\frac{\partial}{\partial z'}.
\end{aligned}
\tag{8.3}
$$

Then, for instance, instead of employing such a passivity condition as

$$
\frac{\partial}{\partial z'}L_z=\frac{\partial}{\partial x}L_{z'},
\tag{8.4}
$$

we use

$$
\frac{\partial}{\partial z'}\frac{dL}{dx}-\frac{d}{dx}L_{z'}=L_z+F_{z'}L_{y'}+G_{z'}L_{z'}
\tag{8.5}
$$

and likewise in all similar cases.

We now adopt the following system of cotes for the independent variables and unknown functions of $\mathfrak{S}$:

(8.6)

x	y	z	y'	z'	L	M	N
1	1	1	1	1	0	0	0
0	0	0	0	0	3	1	2
1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0

This is equivalent to arranging all the derivatives of L, M, N first with regard to their total order in all the independent variables, then with regard to the unknown function in the order L, N, M (note that M is last), then as to the order in x, in y, in z, in y', in z'. The cote of x is thought of as associated with the operator d/dx rather than $\partial/\partial x$.

With these cotes, the system $\mathfrak{S}$ is seen to be orthonomic, for it has the three characteristic properties stated in §6.

There are two passivity tests to be applied, the first associated with (8.5), the second with

$$\frac{\partial}{\partial y'}\frac{dN}{dx} - \frac{d}{dx}N_{y'} = N_y + F_{y'}N_{y'} + G_{y'}N_{z'}. \tag{8.7}$$

In all such passivity tests involving the operation d/dx we may proceed systematically as follows. The first three equations of (7.5) represent d/dx of L, M, N as linear expressions in L, M, N. By use of the alternating relations (8.3), we may then represent d/dx of any first partial derivative of L, M, N as a linear expression in these partial derivatives and L, M, N themselves, that is, as a linear differential expression of the first order in L, M, N. Continuing step-by-step in this way, we can represent d/dx of any kth order partial derivative of L, M, N as a linear differential expression of kth order in L, M, N.

Accordingly, by applying d/dx to any linear differential relation in L, M, N we obtain another one *of the same order*. By eliminating from this all principal derivatives, as we may do, we have a relation among parametric derivatives which is a consequence of the original differential relation. According to the general theory of §6, this new relation must be satisfied identically for a passivity condition; otherwise, it must be adjoined to the given differential system so as to form an extended system.

Let us then apply d/dx to the fifth equation (7.5), obtaining

$$\frac{d}{dx}L_{z'} = \frac{d}{dx}M_{y'}.$$

By (8.5) and the first equation of (7.5),

$$\frac{d}{dx}L_{z'} = \frac{\partial}{\partial z'}(-F_{y'}L - G_{y'}M) - L_z - F_{z'}L_{y'} - G_{z'}L_{z'}.$$

Similarly,

$$\frac{d}{dx}M_{y'} = \frac{\partial}{\partial y'}\left[-\tfrac{1}{2}F_{z'}L - \tfrac{1}{2}(F_{y'} + G_{z'})M - \tfrac{1}{2}G_{y'}N\right] - M_y - F_{y'}M_{y'} - G_{y'}M_{z'}.$$

Equating the last two expressions, we find, after writing $L_{z'} = M_{y'}$, $N_{y'} = M_{z'}$, the following relation among parametric derivatives of $\mathfrak{S}$:

$$\begin{aligned} L_z = {} & M_y - \tfrac{1}{2}F_{z'}L_{y'} + \tfrac{1}{2}(F_{y'} - G_{z'})M_{y'} + \tfrac{1}{2}G_{y'}M_{z'} \\ & - \tfrac{1}{2}F_{y'z'}L + \tfrac{1}{2}(F_{y'y'} - G_{y'z'})M + \tfrac{1}{2}G_{y'y'}N. \end{aligned} \tag{8.8}$$

Treating the sixth equation of (7.5) in a similar way:

$$\frac{d}{dx}N_{y'} = \frac{d}{dx}M_{z'},$$

we get

$$
\begin{aligned}
N_y = M_z &+ \tfrac{1}{2}F_{z'}M_{y'} - \tfrac{1}{2}(F_{y'} - G_{z'})M_{z'} - \tfrac{1}{2}G_{y'}N_{z'} \\
&+ \tfrac{1}{2}F_{z'z'}L - \tfrac{1}{2}(F_{y'z'} - G_{z'z'})M - \tfrac{1}{2}G_{y'z'}N.
\end{aligned} \tag{8.9}
$$

The relations (8.8), (8.9) must be adjoined to (7.5), or $\mathfrak{S}$, so giving the first extended system $\mathfrak{S}_1$, which is seen to be arranged in orthonomic form.

We have next to form the passivity conditions of $\mathfrak{S}_1$.

First we apply d/dx to (8.8) and (8.9). We express all such quantities as d/dx of L_z, M_y, $L_{y'}$, etc., in the manner previously described and then eliminate all principal derivatives, taking due account of the defining conditions of Case I:

$$A = 0, \qquad B = 0, \qquad C = 0.$$

Then the resulting relation among parametric derivatives is found to *disappear identically*. Thus, these particular passivity conditions are satisfied, and no new equations need be adjoined on their account.

There are two further passivity conditions of $\mathfrak{S}_1$ to consider, represented by

$$\frac{\partial L_z}{\partial z'} = \frac{\partial L_{z'}}{\partial z}, \qquad \frac{\partial N_y}{\partial y'} = \frac{\partial N_{y'}}{\partial y}.$$

Upon calculation, *these two conditions turn out to be the same*, namely:

$$
\begin{aligned}
M_{yz'} = M_{zy'} &+ \tfrac{1}{2}F_{z'}M_{y'y'} - \tfrac{1}{2}(F_{y'} - G_{z'})M_{y'z'} - \tfrac{1}{2}G_{y'}M_{z'z'} \\
&+ \tfrac{1}{2}F_{z'z'}L_{y'} + \tfrac{1}{2}G_{z'z'}M_{y'} - \tfrac{1}{2}F_{y'y'}M_{z'} - \tfrac{1}{2}G_{y'y'}N_{z'} \\
&+ \tfrac{1}{2}F_{y'z'z'}L - \tfrac{1}{2}(F_{y'y'z'} - G_{y'z'z'})M - \tfrac{1}{2}G_{y'y'z'}N.
\end{aligned} \tag{8.10}
$$

The relation (8.10) among parametric derivatives of $\mathfrak{S}_1$ must be adjoined to that system to form a second extended system $\mathfrak{S}_2$.

The system $\mathfrak{S}_2$ is subject to just one passivity test, namely that in which we take d/dx of (8.10) and then substitute for d/dx of $M_{yz'}$, $M_{zy'}$, $M_{y'y'}$, etc., the linear differential expressions of second order obtainable in the manner previously described—after which we eliminate all principal derivatives. When this is done, with due account taken of $A=0$, $B=0$, $C=0$, it is found after long calculations that the result is an identity: $0=0$.

Therefore: *in Case* I, *the second extension* $\mathfrak{S}_2$ *of the differential system* $\mathfrak{S}$ *is passive.*

$\mathfrak{S}_2$, we recall, consists of (7.5)—except $AL+BM+CN=0$, which disappears identically—(8.8), (8.9), (8.10).

$\mathfrak{S}_2$ is also orthonomic, as we see by reference to the table of cotes (8.6); therefore, by the Riquier theory, $\mathfrak{S}_2$ is completely integrable.

A complete set of parametric derivatives of $\mathfrak{S}_2$ with their corresponding multipliers is seen, by the rules of the theory of differential systems(22), to be

(22) Cf. J. M. Thomas, loc. cit., §§75, 76.

$$
\begin{aligned}
&L\text{: multipliers } y,\ y';\\
&N\text{: multipliers } z,\ z';\\
&M\text{: multipliers } y,\ z,\ y';\\
&M_{s'}\text{: multipliers } z,\ y',\ z'.
\end{aligned}
\tag{8.11}
$$

This means that every parametric derivative of $\mathfrak{S}_2$ can be expressed in one and only one of the forms: $\partial^{r+s}/\partial y^r \partial y'^s$ of L, $\partial^{r+s}/\partial z^r \partial z'^s$ of N, $\partial^{r+s+t}/\partial y^r \partial z^s \partial y'^t$ of M, $\partial^{r+s+t}/\partial z^r \partial y'^s \partial z'^t$ of $M_{s'}$.

It follows that a solution of $\mathfrak{S}_2$ (equivalent to $\mathfrak{S}$) exists and is uniquely determined if we assign as arbitrary analytic functions, regular about (0, 0, 0, 0, 0) (which, without essential loss of generality, we may take as initial values of x, y, z, y', z'), the following:

$$
\begin{aligned}
f(y, y') &= L(0, y, 0, y', 0),\\
g(z, z') &= N(0, 0, z, 0, z'),\\
h(y, z, y') &= M(0, y, z, y', 0),\\
k(z, y', z') &= M_{s'}(0, 0, z, y', z').
\end{aligned}
\tag{8.12}
$$

It is only necessary to provide that

$$f(0, 0)g(0, 0) - h(0, 0, 0)^2 \neq 0 \tag{8.13}$$

in order to take care of the inequation $LN - M^2 \neq 0$.

Instead of (0, 0, 0, 0, 0), we may use any system of initial values $(x_0, y_0, z_0, y_0', z_0')$ about which functions F, G are regular analytic.

Thus, the general solution of $\mathfrak{S}$ in Case I depends on *two* arbitrary functions of *three* arguments, and *two* arbitrary functions of *two* arguments. As will be seen by comparison with the later cases, this is the highest possible degree of generality of the variation problem corresponding to a given curve family as extremals.

The results of this section have been summarized in Theorem I, stated in §3.

Example. Case I is illustrated in a fairly general way by the example

$$y'' = f(z'), \qquad z'' = 0, \tag{8.14}$$

f denoting an arbitrary function; this is easily verified to obey the requisite conditions (8.2). For $f(z') = 0$, we have the particular case of the straight lines, already treated in the paper by G. Hamel cited in §2.

We may use this example to show also the advantages that may be derived from a knowledge of the finite equations of the given curve family.

The differential equations (8.14) may be integrated, giving

$$y = \tfrac{1}{2}f(a_1)x^2 + a_3x + a_4, \qquad z = a_1x + a_2, \tag{8.15}$$

which expresses the given curve family in terms of four parameters: a_1, a_2, a_3, a_4. By differentiation of (8.15),

$$y' = f(a_1)x + a_3, \qquad z' = a_1. \tag{8.16}$$

The system (8.15), (8.16) may be solved for a_1, a_2, a_3, a_4:

$$\begin{aligned} a_1 &= z', \\ a_2 &= z - z'x, \\ a_3 &= y' - f(z')x, \\ a_4 &= y - y'x + \tfrac{1}{2}f(z')x^2; \end{aligned} \tag{8.17}$$

and we may transform from x, y, z, y', z' as independent variables to x, a_1, a_2, a_3, a_4. d/dx then denotes an operation of differentiation as to x while a_1, a_2, a_3, a_4 are kept constant, that is, partial differentiation as to x of a function of x, a_1, a_2, a_3, a_4.

The first three equations of (7.5) now become

$$\frac{dL}{dx} = 0, \qquad \frac{dM}{dx} = -\tfrac{1}{2}f'(a_1)L, \qquad \frac{dN}{dx} = -f'(a_1)M. \tag{8.18}$$

This is a linear system of ordinary differential equations, whose general solution is readily expressed by means of three parameters λ, μ, ν, arbitrary functions of a_1, a_2, a_3, a_4, as follows:

$$\begin{aligned} L &= \lambda, \\ M &= -\tfrac{1}{2}f'(a_1)\lambda x + \mu, \\ N &= \tfrac{1}{4}f'^2(a_1)\lambda x^2 - f'(a_1)\mu x + \nu. \end{aligned} \tag{8.19}$$

We substitute these values in the other two relations of (7.5): $L_{z'} = M_{y'}$, $N_{y'} = M_{z'}$, and transform the independent variables to x, a_1, a_2, a_3, a_4 according to (8.17); in this way we arrive at the following differential system, where the subscripts 1, 2, 3, 4 denote differentiation as to a_1, a_2, a_3, a_4, respectively:

$$\begin{aligned} \lambda_1 &= \mu_3, \\ \lambda_2 &= \mu_4 - \tfrac{1}{2}f'(a_1)\lambda_3, \\ \nu_3 &= \mu_1, \\ \nu_4 &= \mu_2 + \tfrac{1}{2}f'(a_1)\mu_3 + \tfrac{1}{2}f''(a_1)\lambda. \end{aligned} \tag{8.20}$$

This replaces completely the original differential system (7.5).

The system (8.20) has two passivity conditions: $\lambda_{12} = \lambda_{21}$, $\nu_{34} = \nu_{43}$; but on calculation these give the same relation among parametric derivatives of (8.20) (as was to be expected from our previous general theory; cf. the statement preceding (8.10)), namely:

$$\mu_{14} = \mu_{23} + \tfrac{1}{2}f'(a_1)\mu_{33} + \tfrac{1}{2}f''(a_1)\lambda_3. \tag{8.21}$$

Since there is no other equation in the system (8.20), (8.21) whose first member is a derivative of μ, there is no further passivity test to be made, and the system in question is *passive*.

In the case of the straight lines, $f(a_1)=0$, and the passive system to be integrated reduces to

$$\lambda_1 = \mu_3, \quad \lambda_2 = \mu_4; \qquad \nu_3 = \mu_1, \quad \nu_4 = \mu_2; \qquad \mu_{14} = \mu_{23}. \tag{8.22}$$

A general solution of the last equation, in μ alone, is uniquely determined if $\mu(a_1, a_2, a_3, 0)$, $\mu_4(0, a_2, a_3, a_4)$ are assigned arbitrarily. Then the differentials $\mu_3 da_1 + \mu_4 da_2$ and $\mu_1 da_3 + \mu_2 da_4$ are exact, and give by their integration:

$$\lambda = \int (\mu_3 da_1 + \mu_4 da_2) + g(a_3, a_4),$$

$$\nu = \int (\mu_1 da_3 + \mu_2 da_4) + h(a_1, a_2),$$

where the functions g, h are arbitrary. By (8.19) with $f(a_1)=0$, $(L, M, N) = (\lambda, \mu, \nu)$, where, by (8.17), we may transform back to x, y, z, y', z' as independent variables. Then ϕ can be found by the procedure of §5.

9. **Case II. The critical cone.** Case II, the most complicated and varied in its results and treatment, is that of rank 1 of the matrix Δ, expressed by

$$\left\| \begin{matrix} A & B & C \\ A_1 & B_1 & C_1 \end{matrix} \right\| = 0, \qquad \|A \; B \; C\| \neq 0. \tag{9.1}$$

These conditions imply the existence of a factor $r(x, y, z, y', z')$ such that

$$A_1 = rA, \qquad B_1 = rB, \qquad C_1 = rC. \tag{9.2}$$

Then by the recursion formulas (7.9), we have, with a common factor s, $A_2 = sA$, $B_2 = sB$, $C_2 = sC$, so that (9.1) indeed expresses rank 1 of the matrix Δ.

Let us consider the equation

$$LN - M^2 = 0, \tag{9.3}$$

whose negation constitutes the basic inequation of the system (7.5) or $\mathfrak{S}$. If L, M, N are interpreted as cartesian coordinates in an auxiliary 3-space, this equation represents a quadric cone with vertex at the origin. This cone, which will play a very important part in our theory, will be referred to as the "*critical cone*" K.

It is evident that an arbitrary point on the critical cone is represented in terms of parameters ρ, ξ by

$$L = \rho\xi^2, \qquad M = \rho\xi, \qquad N = \rho. \tag{9.4}$$

We may also say that these equations represent any vector from the origin along an element of the cone; ξ fixes the element and ρ determines the position of the end-point of the vector on that element.

The linear algebraic equation (7.5_4) of the system $\mathfrak{S}$:

$$AL + BM + CN = 0 \tag{9.5}$$

—not disappearing identically because of the second condition (9.1)—represents a *plane* $\mathcal{P}$ in the space (L, M, N) passing through the vertex of the critical cone. A definite position of this plane corresponds to any given values of (x, y, z, y', z') considered as parameters.

The intersection of the plane $\mathcal{P}$ with the critical cone K is determined by substituting (9.4) in (9.5), which leads to the quadratic equation

$$A\xi^2 + B\xi + C = 0. \tag{9.6}$$

The two values of ξ furnished by this equation determine the two elements in which the plane $\mathcal{P}$ intersects the critical cone. These elements may be real and distinct, conjugate imaginary, or real and coincident.

We shall denote by λ, μ the two roots of the fundamental quadratic equation (9.6)([23]); so that we have

$$A\lambda^2 + B\lambda + C = 0, \qquad A\mu^2 + B\mu + C = 0; \tag{9.7}$$

$$\lambda = \frac{-B + (B^2 - 4AC)^{1/2}}{2A}, \qquad \mu = \frac{-B - (B^2 - 4AC)^{1/2}}{2A}; \tag{9.8}$$

$$\lambda + \mu = -B/A, \qquad \lambda\mu = C/A. \tag{9.9}$$

It is of fundamental importance for the subsequent theory to know whether these roots λ, μ—known functions of x, y, z, y', z' determined by the given curve family—are *distinct* or *coincident*, that is, whether the plane $\mathcal{P}$ intersects the critical cone in two distinct elements (real or conjugate imaginary) or, on the other hand, is tangent to this cone along an element.

We therefore divide the following discussion into two cases accordingly:

Case IIa. $B^2 - 4AC \neq 0$, $\lambda \neq \mu$;

Case IIb. $B^2 - 4AC = 0$, $\lambda = \mu$.

10. **Case IIa. The plane $\mathcal{P}$ intersects the critical cone K in two distinct elements.** Let e_1, e_2 denote the two distinct elements of intersection; then these may be considered as coordinate axes in the plane $\mathcal{P}$. According to

([23]) If $A = 0$, one of λ, μ is infinite; and if $A = 0$, $B = 0$, both λ and μ are infinite. These difficulties may be met, provided $C \neq 0$, by interchanging the y and z coordinates, which also interchanges F and G, and will be seen by (7.6) to have the effect of interchanging A and $-C$, while converting B into $-B$. There remains only the case $A = 0$, $C = 0$, $B \neq 0$ (if also $B = 0$, we are in Case I). By the defining conditions (9.1) of Case II, this implies $A_1 = 0$, $C_1 = 0$, and, by reference to (7.6), (7.9): $F_{z'} = 0$, $F_z = 0$, $G_{y'} = 0$, $G_y = 0$. We thus have precisely the example (11.5), which is given separate treatment later on.

(9.4), arbitrary vectors from the origin along e_1, e_2 respectively are represented by

$$L = \rho\lambda^2, \quad M = \rho\lambda, \quad N = \rho;$$
$$L = \sigma\mu^2, \quad M = \sigma\mu, \quad N = \sigma;$$

and any vector from the origin lying in the plane $\mathcal{P}$ is representable as the sum of two such vectors:

$$L = \rho\lambda^2 + \sigma\mu^2, \quad M = \rho\lambda + \sigma\mu, \quad N = \rho + \sigma. \tag{10.1}$$

Here ρ, σ may be regarded as the oblique cartesian coordinates (real or conjugate imaginary) of any point (L, M, N) in the plane $\mathcal{P}$, referred to the vectors $(\lambda^2, \lambda, 1)$, $(\mu^2, \mu, 1)$ along the elements e_1, e_2 respectively as unit vectors.

Our next step is to make the substitution (10.1) in the equations of the system $\mathfrak{S}$ or (7.5), so deriving a differential system in the two unknown functions $\rho(x, y, z, y', z')$, $\sigma(x, y, z, y', z')$ that will be fundamental in the treatment of Case IIa. Before proceeding to do this, we establish the following relations:

$$\begin{aligned} \frac{d\lambda}{dx} &= \tfrac{1}{2}F_{z'}\lambda^2 - \tfrac{1}{2}(F_{y'} - G_{z'})\lambda - \tfrac{1}{2}G_{y'}, \\ \frac{d\mu}{dx} &= \tfrac{1}{2}F_{z'}\mu^2 - \tfrac{1}{2}(F_{y'} - G_{z'})\mu - \tfrac{1}{2}G_{y'}. \end{aligned} \tag{10.2}$$

Their proof is as follows. By differentiation of (9.7),

$$(2A\lambda + B)\frac{d\lambda}{dx} + \left(\lambda^2\frac{dA}{dx} + \lambda\frac{dB}{dx} + \frac{dC}{dx}\right) = 0. \tag{10.3}$$

By (9.2) and the formulas (7.9),

$$\begin{aligned} \frac{dA}{dx} &= (F_{y'} + r)A + \tfrac{1}{2}F_{z'}B, \\ \frac{dB}{dx} &= G_{y'}A + (\tfrac{1}{2}F_{y'} + \tfrac{1}{2}G_{z'} + r)B + F_{z'}C, \\ \frac{dC}{dx} &= \tfrac{1}{2}G_{y'}B + (G_{z'} + r)C. \end{aligned} \tag{10.4}$$

Substituting (10.4) in (10.3), we get, with account taken of (9.7):

$$\begin{aligned} (2A\lambda + B)\frac{d\lambda}{dx} + \tfrac{1}{2}F_{y'}\lambda(2A\lambda + B) + \tfrac{1}{2}G_{y'}(2A\lambda + B) \\ + \tfrac{1}{2}G_{z'}(B\lambda + 2C) + \tfrac{1}{2}F_{z'}\lambda(B\lambda + C) = 0. \end{aligned} \tag{10.5}$$

Now by (9.8).

$$2A\lambda + B = (B^2 - 4AC)^{1/2};$$

and by (9.7, 8),

$$B\lambda + 2C = -\lambda(2A\lambda + B) = -\lambda(B^2 - 4AC)^{1/2}.$$

Substituting in (10.5) and cancelling the factor $(B^2-4AC)^{1/2}\neq 0$, we obtain (10.2), as was desired.

Let us then write (10.1) in the first three equations of (7.5); these become, with the help of (10.2):

$$\begin{aligned}
&\lambda^2\left\{\frac{d\rho}{dx} + (F_{x'}\lambda + G_{x'})\rho\right\} + \mu^2\left\{\frac{d\sigma}{dx} + (F_{x'}\mu + G_{x'})\sigma\right\} = 0,\\
&\lambda\left\{\frac{d\rho}{dx} + (F_{x'}\lambda + G_{x'})\rho\right\} + \mu\left\{\frac{d\sigma}{dx} + (F_{x'}\mu + G_{x'})\sigma\right\} = 0, \qquad (10.6)\\
&\left\{\frac{d\rho}{dx} + (F_{x'}\lambda + G_{x'})\rho\right\} + \left\{\frac{d\sigma}{dx} + (F_{x'}\mu + G_{x'})\sigma\right\} = 0.
\end{aligned}$$

Since $\lambda-\mu\neq 0$, these equations obviously imply the zero value of each of the two brackets, that is:

$$\frac{d\rho}{dx} = -(F_{x'}\lambda + G_{x'})\rho, \qquad \frac{d\sigma}{dx} = -(F_{x'}\mu + G_{x'})\sigma. \qquad (10.7)$$

The fourth equation, $AL+BM+CN=0$, of (7.5) disappears identically after the substitution (10.1).

The remaining equations of (7.5),

$$L_{x'} = M_{y'}, \qquad N_{y'} = M_{x'},$$

give, by the substitution (10.1), two relations, which may be solved for $\rho_{y'}$, $\sigma_{y'}$; we thus obtain

$$\rho_{y'} = \lambda\rho_{x'} + (\lambda_{x'} + \alpha)\rho + \beta\sigma, \qquad \sigma_{y'} = \mu\sigma_{x'} + (\mu_{x'} - \beta)\sigma - \alpha\rho, \qquad (10.8)$$

where

$$\alpha = \frac{\lambda\lambda_{x'} - \lambda_{y'}}{\lambda - \mu}, \qquad \beta = \frac{\mu\mu_{x'} - \mu_{y'}}{\lambda - \mu}. \qquad (10.9)$$

The inequation $LN-M^2\neq 0$ of the system $\mathfrak{S}$ means that the point (L, M, N) shall not lie upon the critical cone; according to the remarks connected with (10.1), this is expressed by the requirement that both

$$\rho \neq 0, \qquad \sigma \neq 0. \qquad (10.10)$$

The same is also seen from the formula: $LN-M^2=\rho\sigma(\lambda-\mu)^2$, following from (10.1)—since $\lambda-\mu\neq 0$.

The fundamental differential system $\mathfrak{T}$ *of Case* IIa *consists of the four equations* (10.7), (10.8) *and the inequations* (10.10).

The vanishing or nonvanishing of the quantities α, β defined by (10.9) is of fundamental importance in the study of the differential system $\mathfrak{T}$; accordingly, we make the following division of Case IIa into subcases:

Case IIa1. The "*separated case*":

$$\lambda\lambda_{x'} - \lambda_{y'} = 0, \qquad \mu\mu_{x'} - \mu_{y'} = 0. \tag{10.11}$$

Case IIa2. The "*semi-separated case*":

$$\lambda\lambda_{x'} - \lambda_{y'} = 0, \qquad \mu\mu_{x'} - \mu_{y'} \neq 0. \tag{10.12}$$

Case IIa3. The "*non-separated case*":

$$\lambda\lambda_{x'} - \lambda_{y'} \neq 0, \qquad \mu\mu_{x'} - \mu_{y'} \neq 0. \tag{10.13}$$

Since λ, μ are defined as the roots of $A\xi^2+B\xi+C=0$, the preceding conditions can also be expressed in terms of A, B, C; we find, by calculations that the reader may easily supply, for the separated case:

$$\begin{aligned} I &\equiv B(BA_{x'} - AB_{x'}) + A(BA_{y'} - AB_{y'}) + A(AC_{x'} - CA_{x'}) = 0, \\ J &\equiv B(BC_{y'} - CB_{y'}) + C(BC_{x'} - CB_{x'}) + C(CA_{y'} - AC_{y'}) = 0; \end{aligned} \tag{10.14}$$

for the semi-separated case:

$$C^2I^2 + (B^3 - 3ABC)IJ + A^2J^2 = 0, \qquad \text{with either } I \neq 0, \text{ or } J \neq 0; \tag{10.15}$$

for the non-separated case:

$$C^2I^2 + (B^3 - 3ABC)IJ + A^2J^2 \neq 0. \tag{10.16}$$

The names we have given to our three subcases refer to the fact that, first, when $\alpha=0, \beta=0$, the system $\mathfrak{T}$ separates into two parts involving respectively ρ alone and σ alone:

$$\frac{d\rho}{dx} = -(F_{x'}\lambda + G_{x'})\rho, \qquad \rho_{y'} = \lambda\rho_{x'} + \lambda_{x'}\rho; \tag{10.17}$$

and

$$\frac{d\sigma}{dx} = -(F_{x'}\mu + G_{x'})\sigma, \qquad \sigma_{y'} = \mu\sigma_{x'} + \mu_{x'}\sigma. \tag{10.18}$$

Second, when $\alpha=0, \beta\neq 0$, we have a system of two equations involving σ alone, namely:

$$\frac{d\sigma}{dx} = -(F_{x'}\mu + G_{x'})\sigma, \qquad \sigma_{y'} = \mu\sigma_{x'} + (\mu_{x'} - \beta)\sigma, \tag{10.19}$$

combined with a mixed system where σ and ρ occur together:

$$(10.20) \qquad \frac{d\rho}{dx} = -(F_{z'}\lambda + G_{z'})\rho, \qquad \rho_{y'} = \lambda\rho_{z'} + \lambda_{z'}\rho + \beta\sigma;$$

hence the name "semi-separated."

Finally, when $\alpha \neq 0$, $\beta \neq 0$, we have the original system (10.7), (10.8) with no separation into partial systems one or both of which involve only one of the unknowns ρ, σ; hence "non-separated."

Throughout the discussion of Case IIa, in all its parts, we shall adopt the following system of cotes:

(10.21)

x	y	y'	z	z'	ρ	σ
1	1	1	1	1	0	0
1	0	0	0	0	0	0
0	1	0	0	0	0	0
0	0	1	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	1	0	0
0	0	0	0	0	2	1

This is equivalent to arranging the derivatives of ρ, σ first as to their total order, then as to their order in x, in y, in y', in z, in z', then as to the unknown function in the order ρ, σ.

We shall solve all our equations for the derivative of highest cote which appears effectively, or—even if this is not done explicitly—our equations may always be thought of as solved for this highest derivative. Thus all the differential systems to be treated will have the *orthonomic* form essential for application of the Riquier theory. Sometimes, however, a form in which some additional terms are transposed will be employed as technically more desirable. This will be done on those occasions when we shall form the passivity condition of two equations by building the alternant of corresponding linear differential operators of first order. Advantageous here will be the simplification which arises from the fact that the alternant is again of the first order, all second order terms which are apparently introduced cancelling. This device of alternants will not interfere, however, with the applicability of the Riquier theory of orthonomic systems, in which it produces only inessential modifications.

11. **Case IIa1. The separated case.** Here we have to solve the system consisting of (10.17), (10.18), (10.10), under the conditions (10.11).

Each of the partial systems comes under the classic theory of simultaneous linear partial differential equations of first order in a single unknown func-

tion(24). Forming the alternant of the two equations of (10.17), we find

$$\rho_y = \lambda\rho_x + H\rho_{x'} + P_1\rho, \tag{11.1}$$

where

$$H = \frac{d\lambda}{dx} = \tfrac{1}{2}F_{x'}\lambda^2 - \tfrac{1}{2}(F_{y'} - G_{x'})\lambda - \tfrac{1}{2}G_{y'}, \qquad \text{by (10.2)},$$

$$P_1 = \lambda_x + H_{x'}.$$

If we next form the alternant of (11.1) with (10.17_1), this can be calculated to disappear identically as a consequence of the equations already in the system.

The alternant of (11.1) with (10.16_2) can be calculated to be

$$(\lambda\lambda_{x'} - \lambda_{y'})\rho_x + U\rho_{x'} + X_1\rho = 0, \tag{11.2}$$

where

$$U = \frac{d}{dx}(\lambda\lambda_{x'} - \lambda_{y'}) + (F_{y'} - \lambda F_{x'})(\lambda\lambda_{x'} - \lambda_{y'}),$$

$$X_1 = U_{x'} + \frac{\partial}{\partial z}(\lambda\lambda_{x'} - \lambda_{y'}).$$

But since $\lambda\lambda_{x'} - \lambda_{y'} = 0$, equation (11.2) *disappears identically*; therefore (10.17), (11.1) constitute a *complete system* for ρ.

Similarly, we find the following complete system for σ: (10.18) together with

$$\sigma_y = \mu\sigma_x + K\sigma_{x'} + Q_1\sigma, \tag{11.3}$$

where

$$K = \frac{d\mu}{dx} = \tfrac{1}{2}F_{x'}\mu^2 - \tfrac{1}{2}(F_{y'} - G_{x'})\mu - \tfrac{1}{2}G_{y'}, \qquad \text{by (10.2)},$$

$$Q_1 = \mu_x + K_{x'}.$$

Indeed, the alternant of (11.3) with (10.18_2) is

$$(\mu\mu_{x'} - \mu_{y'})\sigma_x + V\sigma_{x'} + Y_1\sigma = 0, \tag{11.4}$$

where

$$V = \frac{d}{dx}(\mu\mu_{x'} - \mu_{y'}) + (F_{y'} - \mu F_{x'})(\mu\mu_{x'} - \mu_{y'}),$$

$$Y_1 = V_{x'} + \frac{\partial}{\partial z}(\mu\mu_{x'} - \mu_{y'});$$

(24) For a presentation of this theory, see Goursat-Hedrick, *Differential Equations* (vol. 2, part 2 of *A Course in Mathematical Analysis*, Boston, 1917), p. 265 et seq.

so that the equation (11.4) disappears identically because of $\mu\mu_{z'}-\mu_{y'}=0$.

Since each function ρ, σ involves *five* independent arguments and is a solution of a complete system of *three* equations, the corresponding variation problem has the generality ∞^{2}[(2)]([25]).

Example.

$$y''=F(x, y, y'), \qquad z''=G(x, z, z'), \qquad B\neq 0. \tag{11.5}$$

As we shall see, our inverse problem separates into one in the xy-plane and one in the xz-plane.

It is evident by (7.6) and (7.9) that $A=0$, $C=0$, $A_1=0$, $C_1=0$, so that by (9.1) we are indeed in Case II.

Since by hypothesis, $B\neq 0$, the formula $AL+BM+CN=0$ gives $M=0$ as the equation of the fundamental plane $\mathcal{P}$. We may regard L, N instead of ρ, σ as the coordinates in this plane. The system (7.5) becomes in the present case:

$$\frac{dL}{dx}+F_{y'}L=0, \qquad L_{z'}=0; \tag{11.6}$$

$$\frac{dN}{dx}+G_{z'}N=0, \qquad N_{y'}=0; \tag{11.7}$$

which is clearly of the separated type. Forming alternants, we find

$$L_z=0, \qquad N_y=0. \tag{11.8}$$

By (11.8), and the second equations of (11.6), (11.7),

$$L=L(x, y, y'), \qquad N=N(x, z, z'). \tag{11.9}$$

Equations (11.6_1), (11.7_1) may then be written

$$\begin{aligned}\left[\frac{\partial}{\partial x}+y'\frac{\partial}{\partial y}+F(x, y, y')\frac{\partial}{\partial y'}\right]L&=-F_{y'}L,\\ \left[\frac{\partial}{\partial x}+z'\frac{\partial}{\partial z}+G(x, z, z')\frac{\partial}{\partial z'}\right]N&=-G_{z'}N.\end{aligned} \tag{11.10}$$

Since $L=\phi_{y'y'}$, $M=\phi_{y'z'}$, $N=\phi_{z'z'}$, the condition $M=0$ together with (11.9) implies for ϕ the separated form

$$\phi=\theta(x, y, y')+\psi(x, z, z'), \tag{11.11}$$

within an additive expression of the linear form

$$\alpha(x, y, z)+\beta(x, y, z)y'+\gamma(x, y, z)z'. \tag{11.12}$$

Clearly,

([25]) Cf. Goursat-Hedrick, loc. cit., p. 270.

$$L = \theta_{y'y'}, \qquad N = \psi_{z'z'}. \tag{11.13}$$

The formulas (11.10), (11.13) are precisely those which are met with in solving the following inverse problems in the xy- and xz-planes respectively([25]):

(a) find $\int\theta(x, y, y')dx = \min$, given the extremals $y'' = F(x, y, y')$;

(b) find $\int\psi(x, z, z')dx = \min$, given the extremals $z'' = G(x, z, z')$.

According to (11.11), the solutions of these planar problems are to be added to give the solution of the space problem. The undetermined additive expression (11.12) can be proved, as usual, to be an arbitrary total derivative, $\nu_x + \nu_y y' + \nu_z z'$ where $\nu = \nu(x, y, z)$.

12. **Case IIa2. The semi-separated case.** Here we have to deal with the differential system (10.19), (10.20), (10.10), under the conditions (10.12).

The alternant of the two equations of (10.19) is

$$\sigma_y = \mu\sigma_z + K\sigma_{z'} + Q\sigma, \tag{12.1}$$

while the integrability condition of the two equations of (10.20), arising from formation of the alternant

$$\frac{d}{dx}\left(\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}\right) - \left(\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}\right)\frac{d}{dx}$$

applied to ρ, is

$$\rho_y = \lambda\rho_z + H\rho_{z'} + P_1\rho + S\sigma. \tag{12.2}$$

Here

$$\begin{aligned} H &= \frac{d\lambda}{dx} = \tfrac{1}{2}F_{z'}\lambda^2 - \tfrac{1}{2}(F_{y'} - G_{z'})\lambda - \tfrac{1}{2}G_{y'}, \\ K &= \frac{d\mu}{dx} = \tfrac{1}{2}F_{z'}\mu^2 - \tfrac{1}{2}(F_{y'} - G_{z'})\mu - \tfrac{1}{2}G_{y'}; \\ P_1 &= \lambda_z + H_{z'}, \\ Q &= \mu_z + K_{z'} - S; \\ S &= -\frac{d\beta}{dx} - (F_{y'} - \mu F_{z'})\beta. \end{aligned} \tag{12.3}$$

The alternant of (10.19_1) and (12.1) can be calculated to disappear identically in virtue of the equations of the system, and the same is true of (10.20_1) and (12.2)

The alternant of (10.19_2) and (12.1) is

$$(\mu\mu_{z'} - \mu_{y'})\sigma_z + V\sigma_{z'} + Y_2\sigma = 0, \tag{12.4}$$

where

([25]) Cf. Bolza. loc. cit. (see Footnote 4).

$$V = \frac{d}{dx}(\mu\mu_{z'} - \mu_{y'}) + (F_{y'} - \mu F_{z'})(\mu\mu_{z'} - \mu_{y'}),$$

(12.5)

$$Y_2 = V_{z'} + \frac{\partial}{\partial z}(\mu\mu_{z'} - \mu_{y'}) + S_{y'} - \mu S_{z'} - \beta_y + \mu\beta_z + K\beta_{z'}.$$

It is remarkable that precisely the same equation (12.4) *is also the integrability condition of* (10.20_2) and (12.2), as found by forming the alternant of the linear differential operators

$$\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}$$

and

$$\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - H\frac{\partial}{\partial z'}$$

applied to ρ. It follows that if any solution σ of the system (10.19), (12.1), (12.4) is substituted in (10.20), (12.2), *this becomes a complete system for* ρ.

All depends then on solving the system (10.19), (12.1), (12.4) for σ. It is evident that the formation of further alternants—(12.4) with (10.20_1), with (10.20_2), with (12.1)—gives new equations of the form

$$\cdot\sigma_{z'} + \cdot\sigma = 0, \tag{12.6}$$

where the dots represent calculable known coefficients. There are just two of this type, since the one arising from (10.20_1) and (12.4) can be shown to disappear identically.

If the two equations (12.6) are independent as algebraic linear equations in $\sigma_{z'}$, σ (determinant $\neq 0$), then $\sigma = 0$, in contradiction with our prescribed inequation $\sigma \neq 0$. The given curve family is then nonextremal.

If the two equations (12.6) disappear identically (each coefficient equal to zero), then (10.20_1), (10.20_2), (12.1), (12.4) constitute a complete system of four equations for σ, with a solution, therefore, of the generality $\infty^{1/(1)}$. Any such solution substituted in the system for ρ renders this a *complete* system containing three equations, whose solution therefore has the generality $\infty^{1/(2)}$. The given curve family is consequently of extremal type and the corresponding integrand ϕ has the generality $\infty^{1/(2)+1/(1)}$.

Finally, let the two equations (12.6) amount to exactly one independent equation (coefficient-matrix of rank one). If the coefficient of $\sigma_{z'}$ in this equation is zero, then that of σ cannot be zero (because of the rank one); it follows that $\sigma = 0$, in contradiction with the stipulated inequation $\sigma \neq 0$, and we have a nonextremal family. On the other hand, if the coefficient of $\sigma_{z'}$ is not equal to zero, then we have a system of five equations of the form

$$\sigma_z = \cdot\sigma, \quad \sigma_y = \cdot\sigma, \quad \sigma_z = \cdot\sigma, \quad \sigma_{y'} = \cdot\sigma, \quad \sigma_{z'} = \cdot\sigma, \tag{12.7}$$

where the dots represent known coefficients. The solvability of this system for σ ($\neq 0$) depends on the exactness of the differential

$$\cdot dx + \cdot dy + \cdot dz + \cdot dy' + \cdot dz',$$

where the dots are respectively the coefficients in (12.7). The solution, if existent, is unique up to a constant factor (which may be ignored according to the remarks associated with (3.7)). The generality of the corresponding variation problem is $\infty^{1/(2)}$, arising out of that degree of generality in the system for ρ.

The results of this section have been summarized in Theorem III of §3.

Example. Catenaries. The y-axis in xyz-space being taken vertical, the differential equations

$$y'' = \frac{1 + y'^2 + z'^2}{y}, \qquad z'' = 0 \tag{12.8}$$

represent all vertical catenaries the directrix of each of which coincides with the trace of its plane on the horizontal xz-plane.

Here we calculate by (7.6), (7.9):

$$\begin{aligned} A &= -\frac{4y'z'}{y^2}, \qquad B = -\frac{4(1+z'^2)}{y^2}, \qquad C = 0; \\ A_1 &= \frac{12y'^2z'}{y^3}, \qquad B_1 = \frac{12y'(1+z'^2)}{y^3}, \qquad C_1 = 0. \end{aligned} \tag{12.9}$$

Thus the defining conditions (9.1) of Case II are verified.

The fundamental quadratic equation (9.6) is here

$$y'z'\xi^2 + (1 + z'^2)\xi = 0, \tag{12.10}$$

whose roots are

$$\lambda = 0, \qquad \mu = -\frac{1 + z'^2}{y'z'}. \tag{12.11}$$

The system (10.19, 20) can in this case be calculated to be

$$\frac{d\rho}{dx} = 0, \qquad \rho_{y'} = -\frac{1}{y'z'^2}\sigma; \tag{12.12}$$

$$\frac{d\sigma}{dx} = 2\,\frac{1 + z'^2}{yy'}\sigma, \qquad \sigma_{y'} = -\frac{1 + z'^2}{y'z'}\sigma_{z'} + \frac{2 - z'^2}{y'z'^2}\sigma; \tag{12.13}$$

with the prescribed inequations $\rho \neq 0$, $\sigma \neq 0$.

Forming the alternant of the two equations of (12.12), we get

$$\rho_y = \frac{1 + y'^2 + z'^2}{yy'^2z'^2}\sigma. \tag{12.14}$$

If we form the alternant of the two equations of (12.13), we obtain the equation

$$\sigma_y = -\frac{1+z'^2}{y'z'}\sigma_z + K\sigma_{z'} + Q\sigma, \tag{12.15}$$

where

$$K = \frac{(1+z'^2)(1+y'^2+z'^2)}{yy'^2z'},$$

$$Q = \frac{-2(1+y'^2)+(1+y'^2)z'^2+3z'^4}{yy'^2z'^2}.$$

The alternant of (12.15) with the first of (12.13) disappears identically in virtue of the equations of the system. The alternant of (12.15) with the second of (12.13) is simply

$$\sigma_z = 0. \tag{12.16}$$

(12.13), (12.15), (12.16) can be verified to constitute a complete system for σ, all further alternants being linear consequences of the equations of the system.

The integrability conditions of (12.12), (12.14) are satisfied in virtue of the complete system for σ. Accordingly, all our equations form a completely integrable (passive orthonomic) system in ρ, σ, the generality of whose solution is seen to be $\infty^{1f(2)+1f(1)}$.

13. **Case IIa3. The non-separated case.** It may first be remarked that the class of curve families (F, G) coming under this case is quite extensive. For the conditions for Case IIa3 are that the two partial differential equations of third order in F, G:

$$\frac{A_1}{A} = \frac{B_1}{B} = \frac{C_1}{C}, \tag{13.1}$$

have a solution which obeys at least one of the inequations $A \neq 0$, $B \neq 0$, $C \neq 0$ and also the inequation (10.16). Now, according to standard existence theorems (Cauchy-Kovalevsky), the differential system (13.1), consisting of two equations for two unknown functions, has a solution involving a number of arbitrary functions in its initial data. It is easy to adjust these arbitrary functions so as to obey all the specified inequations.

We proceed, then, to deal with the differential system of Case IIa3, which consists of (10.7), (10.8) and the inequations (10.10), under the conditions (10.13).

Forming the integrability conditions of (10.7_1) with (10.8_1) and of (10.7_2) with (10.8_2), we find

$$\begin{aligned}\rho_y &= \lambda\rho_z + H\rho_{z'} + P\rho + S\sigma,\\ \sigma_y &= \mu\sigma_z + K\sigma_{z'} + Q\sigma + R\rho,\end{aligned} \tag{13.2}$$

where

$$\begin{aligned}H &= \frac{d\lambda}{dx} = \tfrac{1}{2}F_{z'}\lambda^2 - \tfrac{1}{2}(F_{y'} - G_{z'})\lambda - \tfrac{1}{2}G_{y'},\\ K &= \frac{d\mu}{dx} = \tfrac{1}{2}F_{z'}\mu^2 - \tfrac{1}{2}(F_{y'} - G_{z'})\mu - \tfrac{1}{2}G_{y'};\end{aligned}$$

$$\begin{aligned}R &= \frac{d\alpha}{dx} + (F_{y'} - \lambda F_{z'})\alpha,\\ S &= -\frac{d\beta}{dx} - (F_{y'} - \mu F_{z'})\beta;\end{aligned} \tag{13.3}$$

$$\begin{aligned}P &= \lambda_z + H_{z'} - R,\\ Q &= \mu_z + K_{z'} - S.\end{aligned}$$

The integrability condition of (10.7_1) with (13.2_1) disappears identically in virtue of the equations of the system, and the same is true of (10.7_2) and (13.2_2).

On the other hand, the integrability condition of (10.8_1) with (13.2_1) is found after calculation to be

$$(\lambda\lambda_{z'} - \lambda_{y'})\rho_z + U\rho_{z'} + X\rho = (\mu\mu_{z'} - \mu_{y'})\sigma_z + V\sigma_{z'} + Y\sigma, \tag{13.4}$$

where

$$\begin{aligned}U &= \lambda_y - \lambda\lambda_z - H\lambda_{z'} + \lambda H_{z'} - H_{y'}\\ &= \frac{d}{dx}(\lambda\lambda_{z'} - \lambda_{y'}) + (F_{y'} - \lambda F_{z'})(\lambda\lambda_{z'} - \lambda_{y'})\\ &= (H - K)\alpha + (\lambda - \mu)R,\\ V &= \mu_y - \mu\mu_z - K\mu_{z'} + \mu K_{z'} - K_{y'}\\ &= \frac{d}{dx}(\mu\mu_{z'} - \mu_{y'}) + (F_{y'} - \mu F_{z'})(\mu\mu_{z'} - \mu_{y'})\\ &= (H - K)\beta - (\lambda - \mu)S;\end{aligned} \tag{13.5}$$

$$X = \beta R + \alpha S + U_{z'} + \frac{\partial}{\partial z}(\lambda\lambda_{z'} - \lambda_{y'}) + R_{y'} - \lambda R_{z'} + \alpha_y - \lambda\alpha_z - H\alpha_{z'},$$

$$Y = -\beta R - \alpha S + V_{z'} + \frac{\partial}{\partial z}(\mu\mu_{z'} - \mu_{y'}) + S_{y'} - \mu S_{z'} - \beta_y + \mu\beta_z + K\beta_{z'}.$$

If we interchange λ with μ, the effect, as seen by our defining formulas, is to interchange α and $-\beta$, H and K, R and S, P and Q, U and V, X and Y. If at the same time we interchange ρ and σ, the equation (13.4) is seen to go over into itself. This is at the basis of the important fact that *the same equation* (13.4) *is also obtained as the integrability condition of* (10.8_2) *with* (13.2_2).

Let us introduce a third unknown function τ to denote the common value of the two members of (13.4):

$$\begin{aligned} \tau &= (\lambda\lambda_{x'} - \lambda_{y'})\rho_x + U\rho_{x'} + X\rho, \\ \tau &= (\mu\mu_{x'} - \mu_{y'})\sigma_x + V\sigma_{x'} + Y\sigma. \end{aligned} \tag{13.6}$$

Since, by hypothesis, $\lambda\lambda_{x'}-\lambda_{y'}\neq 0$, $\mu\mu_{x'}-\mu_{y'}\neq 0$, we can then give our differential system the form:

$$\begin{aligned} \frac{d\rho}{dx} &= \cdot\rho, \\ \rho_{y'} &= \cdot\rho_{x'} + \cdot\rho + \cdot\sigma, \\ \rho_y &= \cdot\rho_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \rho_x &= \cdot\rho_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau; \\ \frac{d\sigma}{dx} &= \cdot\sigma, \\ \sigma_{y'} &= \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \sigma_y &= \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \sigma_x &= \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau; \end{aligned} \tag{13.7}$$

where the dots represent calculable known coefficients. If now we form the new integrability conditions of this system[27], we find that the resulting equations are solvable for the first partial derivatives of τ (in which circumstance the requirement $\lambda-\mu\neq 0$ plays a part); indeed, we get a result of the form

$$\begin{aligned} \frac{d\tau}{dx} &= \cdot\rho + \cdot\sigma + \cdot\tau\,[28], \\ \tau_y &= \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \tau_x &= \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \tau_{y'} &= \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \tau_{x'} &= \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma + \cdot\tau. \end{aligned} \tag{13.8}$$

[27] We may use the formulas (13.6) as convenient for this purpose, building alternants of the linear operators in the second members of (13.6) with the linear operators occurring in (10.7), (10.8), (13.2).

[28] Two expressions are obtained for $\partial\tau/\partial x$, but these can be calculated to be equal in virtue of our other formulas. The coefficient of τ in the first equation of (13.8) is $-(F_{y'}+G_{x'})$.

The differential system we now have to consider is (13.7), (13.8) in the three unknown functions ρ, σ, τ.

Referring to (10.21), we give τ a last cote of 3 and all other cotes 0, thus ranking τ higher than ρ, σ. The differential system in question then has the orthonomic form.

Forming the integrability conditions of (13.8) (such as $\partial\tau_y/\partial z = \partial\tau_z/\partial y$), we obtain, after reduction with the help of (13.7), a number of equations of the form

$$\cdot\rho_{z'z'} + \cdot\sigma_{z'z'} + \cdot\rho_{z'} + \cdot\sigma_{z'} + \cdot\rho + \cdot\sigma + \cdot\tau = 0. \tag{13.9}$$

If all these equations disappear identically, the system (13.7), (13.8) is passive, and has a solution corresponding to arbitrarily assigned values of $\rho(0, 0, 0, 0, z')$, $\sigma(0, 0, 0, 0, z')$: generality $\infty^{2(1)}$.

Laying this case aside, we may find it possible to eliminate some of the partial derivatives from (13.9) by linear combination of these equations. If the rank of the whole system is r, while the rank of the coefficients of $\rho_{z'z'}$, $\sigma_{z'z'}$, $\rho_{z'}$, $\sigma_{z'}$ is $r-p$ (p necessarily $\geqq 0$), it will be possible to infer in this way exactly p independent relations of the form

$$\cdot\rho + \cdot\sigma + \cdot\tau = 0. \tag{13.10}$$

In case $p=3$, (13.10) implies $\rho=0$, $\sigma=0$, $\tau=0$, in contradiction with the prescribed inequations $\tau\neq 0$, $\sigma\neq 0$.

If $p=2$, it is possible to eliminate τ and derive a relation of the form $\sigma = \cdot\rho$ (or $\rho = \cdot\sigma$). By substitution in the original differential system (10.7), (10.8), this takes the form

$$\rho_z = \cdot\rho, \quad \rho_y = \cdot\rho, \quad \rho_z = \cdot\rho, \quad \rho_{y'} = \cdot\rho, \quad \rho_{z'} = \cdot\rho, \tag{13.11}$$

together with a condition of the form $(\cdot)\rho = 0$—all the dots representing calculable known coefficients. The integrability conditions of (13.11) are of the form $(\cdot)\rho = 0$, so that it is necessary and sufficient for the existence of a solution $\rho\neq 0$ that all the coefficients $(\cdot)$ be equal to zero. If this is the case, ρ is determined uniquely up to a constant factor.

In case only one independent equation of the type (13.10) can be obtained, we may suppose it solvable for τ:

$$\tau = \cdot\rho + \cdot\sigma; \tag{13.12}$$

otherwise we find ourselves in the situation just discussed. Substituting (13.12) in (13.7), we have a system in ρ, σ, whose integrability conditions are those which result by substituting (13.12) in (13.18). The first equation of (13.8) gives in this way a relation of the form $\cdot\rho + \cdot\sigma = 0$ which—if it does not disappear identically, or imply $\rho=0$ or $\sigma=0$—puts us in the preceding case. If this relation, $\cdot\rho + \cdot\sigma = 0$, does disappear identically, the substitution of (13.12) in the other relations of (13.8) may be used, with (13.7), to give in-

tegrability conditions of the form

$$\cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma = 0. \tag{13.13}$$

If (13.13) disappears identically, we have in (13.7) (with τ replaced by its value (13.12)) a passive system in ρ, σ.

If, by linear combination of the equations of (13.13), we can eliminate $\rho_{x'}$, $\sigma_{x'}$, that is, infer a relation of the form $\cdot\rho+\cdot\sigma=0$, we are in a previously discussed case.

If we cannot do this, it must be that the rank of the whole system (13.13) is the same as the rank of the coefficients of $\rho_{x'}$, $\sigma_{x'}$. This common rank may be 2 or 1, since we have disposed of the case of rank zero of the system.

If the rank is 2, we can solve for $\rho_{x'}$, $\sigma_{x'}$:

$$\rho_{x'} = \cdot\rho + \cdot\sigma, \qquad \sigma_{x'} = \cdot\rho + \cdot\sigma. \tag{13.14}$$

Now by substituting (13.12), (13.14) in (13.7), every first partial derivative of ρ, σ is expressed in the form $\cdot\rho+\cdot\sigma$:

$$\begin{aligned}
&\frac{d\rho}{dx} = \cdot\rho, && \frac{d\sigma}{dx} = \cdot\sigma,\\
&\rho_y = \cdot\rho + \cdot\sigma, && \sigma_y = \cdot\rho + \cdot\sigma,\\
&\rho_z = \cdot\rho + \cdot\sigma, && \sigma_z = \cdot\rho + \cdot\sigma,\\
&\rho_{y'} = \cdot\rho + \cdot\sigma, && \sigma_{y'} = \cdot\rho + \cdot\sigma,\\
&\rho_{z'} = \cdot\rho + \cdot\sigma; && \sigma_{z'} = \cdot\rho + \cdot\sigma.
\end{aligned} \tag{13.15}$$

This is a system of a well-known type, easily amenable to treatment by standard methods([29]). It is obvious that here all the integrability conditions are of the form $\cdot\rho+\cdot\sigma=0$. If there are two such independent conditions, we infer $\rho=0$, $\sigma=0$, in violation of the prescribed inequations $\rho\neq0$, $\sigma\neq0$. If there is only one such condition, then, first, neither coefficient may be zero, otherwise we again have contradiction with either $\rho\neq0$ or $\sigma\neq0$. Assuming neither coefficient zero, our system has a solution for ρ, σ, which is essentially unique (within a constant factor). Finally, if all the integrability conditions $\cdot\rho+\cdot\sigma=0$ disappear identically, (13.15) is a completely integrable system, with a solution involving linearly two arbitrary constants (only one of which is essential).

If the rank mentioned in the second preceding paragraph is 1, then (13.13) consists of just one equation solvable for either $\rho_{x'}$ or $\sigma_{x'}$, say the former:

$$\rho_{x'} = \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma. \tag{13.15a}$$

Substituting this, as well as (13.12), in (13.7), we have a differential system in ρ, σ, whose integrability conditions are a set of equations of the form

([29]) Cf. Eisenhart, *An Introduction to Differential Geometry*, Princeton, 1940, p. 114.

(13.15b) $$\cdot\sigma_{s's'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma = 0.$$

If, by elimination, we can derive from these equations a relation of the form $\sigma_{s'} = \cdot\rho + \cdot\sigma$, or $\cdot\rho + \cdot\sigma = 0$ (non-identically vanishing), then we are evidently in previously discussed cases ((13.15) or (13.11)). If not, it must be that (13.15b) contains just one independent equation, and that solvable in the form

(13.15c) $$\sigma_{s's'} = \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma.$$

Adjoining this to the equations already in our possession, we have a differential system in which the only parametric derivatives are $\sigma_{s'}, \rho, \sigma$. All further integrability conditions are therefore of the form

(13.15d) $$\cdot\sigma_{s'} + \cdot\rho + \cdot\sigma = 0.$$

Either all of these disappear identically, in which case we have a passive system, or else we derive at least one relation of the form $\sigma_{s'} = \cdot\rho + \cdot\sigma$, or $\cdot\rho + \cdot\sigma = 0$—wherewith our investigation is reduced, as before, to previously discussed cases.

We have now disposed of the case where even one relation of the form (13.10) can be inferred linearly from (13.9).

If this cannot be done, it may still be possible to eliminate $\rho_{s's'}, \sigma_{s's'}$ from (13.9) and obtain relations of the form

(13.16) $$\cdot\rho_{s'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau = 0.$$

Let r denote the rank of (13.9) and $r-q$ ($q \geqq 0$) the rank of the coefficients of $\rho_{s's'}, \sigma_{s's'}$; then exactly q independent relations of the form (13.16) can be inferred by linear combination of (13.9). We are thus supposing $q \geqq 1$—otherwise no relations of the type (13.16) are derivable by the elimination process mentioned.

If q were as large as 3, we could eliminate $\rho_{s'}, \sigma_{s'}$ from (13.16) and obtain a relation of the form (13.10); but we have supposed this not to be the case.

If $q = 2$, then we can solve the two relations (13.16) for $\rho_{s'}, \sigma_{s'}$ (otherwise it would be possible to eliminate these, contrary to hypothesis) and get:

(13.17) $$\rho_{s'} = \cdot\rho + \cdot\sigma + \cdot\tau, \qquad \sigma_{s'} = \cdot\rho + \cdot\sigma + \cdot\tau.$$

Substituting in (13.7), (13.8), we have a system where every first partial derivative of ρ, σ, τ is expressed in the form $\cdot\rho + \cdot\sigma + \cdot\tau$, that is, a system of the standard type (13.15) with three unknown functions instead of two. All the integrability conditions of this system are of the form $\cdot\rho + \cdot\sigma + \cdot\tau = 0$, so that we have an easy discussion based on the number of such independent conditions (0, 1, or 2) not in contradiction with $\rho \neq 0, \sigma \neq 0$.

If $q = 1$, we can solve the single equation (13.16) for $\rho_{s'}$ or $\sigma_{s'}$, say:

(13.18) $$\rho_{s'} = \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau.$$

The integrability conditions of this with the first group of (13.7) are of the form

$$\cdot\sigma_{s's'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau = 0. \tag{13.19}$$

If we can eliminate $\sigma_{s's'}$ from these equations, we clearly have cases already discussed. Otherwise, (13.19) must consist of just one equation and this solvable for $\sigma_{s's'}$:

$$\sigma_{s's'} = \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau. \tag{13.20}$$

After adjunction of this to the preceding equations of our system, the only partial derivatives which remain are the four: $\sigma_{s'}$, ρ, σ, τ, so that all further integrability conditions must involve only these in linear homogeneous form. Accordingly, with the formation of at most four additional integrability conditions we come to a decision as to the solvability of our system (in conformity with $\rho \neq 0$, $\sigma \neq 0$) and as to the number of arbitrary constants (at most four) in the solution.

Now there remains for the system (13.9) only the case where no relations can be derived by elimination of $\rho_{s's'}$, $\sigma_{s's'}$. In this case, the rank of the entire system is equal to the rank of the coefficients of $\rho_{s's'}$, $\sigma_{s's'}$. The possible values of this common rank s are 2, 1, since $s=0$ has already been laid aside as a case of passivity of the system.

If $s=2$, we can solve (13.19) in the form

$$\begin{aligned} \rho_{s's'} &= \cdot\rho_{s'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau, \\ \sigma_{s's'} &= \cdot\rho_{s'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau. \end{aligned} \tag{13.21}$$

Then the five quantities appearing in the second member are the only remaining parametric derivatives. All further integrability conditions have the form of a linear homogeneous relation in these quantities, so that with the formation of at most five of these relations (supposed independent) we come to a decision.

Finally, let $s=1$. Then (13.9) consists of exactly one relation, which is solvable for either $\rho_{s's'}$ or $\sigma_{s's'}$[30], say:

$$\rho_{s's'} = \cdot\sigma_{s's'} + \cdot\rho_{s'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau. \tag{13.22}$$

The integrability conditions of this with the system (13.7), first part, are of the form

$$\cdot\sigma_{s's's'} + \cdot\sigma_{s's'} + \cdot\rho_{s'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau = 0. \tag{13.23}$$

If we can eliminate $\sigma_{s's's'}$, we are in previously discussed cases. The only alternative is that (13.23) contain just one equation, and this solvable for $\sigma_{s's's'}$:

$$\sigma_{s's's'} = \cdot\sigma_{s's'} + \cdot\rho_{s'} + \cdot\sigma_{s'} + \cdot\rho + \cdot\sigma + \cdot\tau. \tag{13.24}$$

[30] Otherwise, we are obviously in the case associated with (13.16).

After adjunction of this equation, the only parametric derivatives remaining in our differential system are the six which appear in the second member of (13.24). All further integrability conditions are linear homogeneous relations in these quantities, so that after formation of at most six such (independent) relations, we arrive at a decision as to the consistency of our differential system and the degree of generality of its solution.

We may conclude by emphasizing that the determination of the explicit value of any of the coefficients symbolized by dots throughout the preceding discussion requires only differentiations, multiplications, and additions.

We may also call attention to the fact that throughout the discussion of Case IIa, even in the most unfavorable eventualities, the only derivatives of an order higher than the first which have occurred are $\rho_{z'z'}$, $\sigma_{z'z'}$, $\sigma_{z'z'z'}$.

14. **Case IIb. The plane $\mathcal{P}$ is tangent to the critical cone.** In the case indicated by the title of this section, let the element of contact e be the one containing the point $(\lambda^2, \lambda, 1)$. As λ varies, the locus of this point is a conic, whose tangent vector is $(2\lambda, 1, 0)$. The plane $\mathcal{P}$ determined by these two vectors is the one which is tangent to the cone K along the element e. The equation of this plane being $AL+BM+CN=0$, λ is evidently determined as the double root of the quadratic

$$A\lambda^2 + B\lambda + C = 0, \qquad (B^2 - 4AC = 0). \tag{14.1}$$

The coordinates (L, M, N) of an arbitrary point in the plane $\mathcal{P}$ can be written as a linear combination of the two vectors previously mentioned with arbitrary multipliers ρ, σ:

$$L = \lambda^2\rho + 2\lambda\rho, \qquad M = \lambda\rho + \sigma, \qquad N = \rho. \tag{14.2}$$

Here λ is given as a known function of x, y, z, y', z' by (14.1); indeed,

$$\lambda = -\frac{B}{2A} = -\frac{2C}{B}, \tag{14.3}$$

while ρ, σ are unknown functions of x, y, z, y', z'.

Substituting (14.2) in the original differential system $\mathfrak{S}$, or (7.5), we obtain, after reduction, the following differential system for ρ, σ:

$$\begin{aligned} \frac{d\rho}{dx} &= -(F_{z'}\lambda + G_{z'})\rho - F_{z'}\sigma, \\ \frac{d\sigma}{dx} &= -\tfrac{1}{2}(F_{y'} + G_{z'})\sigma, \\ \rho_{y'} &= \lambda\rho_{z'} + \lambda_{z'}\rho + \sigma_{z'}, \\ \sigma_{y'} &= \lambda\sigma_{z'} + (\mathrm{I})\rho + 2\lambda_{z'}\sigma, \end{aligned} \tag{14.4}$$

where

$$(14.5) \qquad (\mathrm{I}) = \lambda\lambda_{x'} - \lambda_{y'}.$$

We must also adjoin the inequation

$$(14.6) \qquad \sigma \neq 0,$$

in order to prevent the point (L, M, N) from lying upon the critical cone.

Our problem in Case IIb is to discuss the solvability for ρ, σ of the differential system consisting of the four equations (14.4) and the inequation (14.6).

To this end, we begin by forming the alternant of (14.4_1), (14.4_3), and also the alternant of (14.4_2), (14.4_4). We find:

$$(14.7) \qquad \begin{aligned} \rho_y &= \lambda\rho_x + (\mathrm{II})\rho_{x'} + (\mathrm{III})\rho + \sigma_x + (\mathrm{IV})\sigma_{x'} + (\mathrm{V})\sigma, \\ \sigma_y &= \lambda\sigma_x + (\mathrm{II})\sigma_{x'} + (\mathrm{VI})\sigma + (\mathrm{VII})\rho, \end{aligned}$$

where the Roman numerals in parentheses denote known functions whose values are:

$$(14.8) \qquad \begin{aligned} (\mathrm{II}) &= \tfrac{1}{2}F_{x'}\lambda^2 - \tfrac{1}{2}(F_{y'} - G_{x'})\lambda - \tfrac{1}{2}G_{y'}, \\ (\mathrm{III}) &= \lambda_x + (\mathrm{II})_{x'}, \\ (\mathrm{IV}) &= F_{x'}\lambda - \tfrac{1}{2}(F_{y'} - G_{x'}), \\ (\mathrm{V}) &= (\mathrm{IV})_{x'}, \\ (\mathrm{VI}) &= 2\lambda_x + 2\lambda_{x'}(\mathrm{IV}) - \lambda(\mathrm{IV})_{x'} + (\mathrm{IV})_{y'} \\ &= 2\lambda_x + F_{x'}(\lambda\lambda_{x'} + \lambda_{y'}) - (F_{y'} - G_{x'})\lambda_{x'} \\ &\quad - F_{x'x'}\lambda^2 + \tfrac{1}{2}(3F_{y'x'} - G_{x'x'})\lambda - \tfrac{1}{2}(F_{y'y'} - G_{y'x'}), \\ (\mathrm{VII}) &= (\mathrm{I})(2F_{x'}\lambda - \tfrac{3}{2}F_{y'} + \tfrac{1}{2}G_{x'}) - \frac{d(\mathrm{I})}{dx} \\ &= \lambda\lambda_x - \lambda_y + (\mathrm{II})\lambda_{x'} - \lambda(\mathrm{II})_{x'} + (\mathrm{II})_{y'} + (\mathrm{I})(\mathrm{IV}). \end{aligned}$$

The equations (14.7) must be adjoined to the original system (14.4). With the same cotes (10.21) that were used in Case IIa, the enlarged system is seen to be in orthonomic form. We proceed to consider this system as to its passivity.

The passivity conditions arising from the equations solved for $d\rho/dx$, $\rho_{y'}$, ρ_y are found, after some calculation, to be *verified identically*. The same is true of the passivity condition arising by comparison of the equations solved for $d\sigma/dx$, σ_y. However, when we apply to the equations solved for $\sigma_{y'}$, σ_y the relation $\partial\sigma_{y'}/\partial y = \partial\sigma_y/\partial y'$—or, what is equivalent, when we form the alternant of the linear differential operators on σ occurring in the two equations—we obtain the passivity condition

$$(14.9) \qquad 2(\mathrm{I})\sigma_x + \{2(\mathrm{I})(\mathrm{IV}) - (\mathrm{VII})\}\sigma_{x'} + (\mathrm{VIII})\rho + (\mathrm{IX})\sigma = 0,$$

where

$$
\begin{aligned}
(\text{VIII}) &= \lambda(\text{VII})_{x'} - (\text{VII})_{y'} + (\text{I})_y - \lambda(\text{I})_z + \lambda_{z'}(\text{VII}) \\
&\quad - (\text{II})(\text{I})_{z'} + (\text{I})(\text{III}) - (\text{I})(\text{VI}), \\
(\text{IX}) &= \lambda(\text{VI})_{x'} - (\text{VI})_{y'} - 2\lambda_{z'z'}(\text{II}) + 2\lambda_{yz'} - 2\lambda\lambda_{zz'} + (\text{I})(\text{V}).
\end{aligned} \tag{14.10}
$$

15. **Case IIb1.** We may now divide the discussion of the differential system (14.4, 6) into two parts as follows:

$$
\begin{aligned}
&\text{Case IIb1.} \quad (\text{I}) \equiv \lambda\lambda_{z'} - \lambda_{y'} = 0, \\
&\text{Case IIb2.} \quad (\text{I}) \equiv \lambda\lambda_{z'} - \lambda_{y'} \neq 0.
\end{aligned} \tag{15.1}
$$

If (I) = 0, the equations (14.4_2), (14.4_4) involve σ alone. Their alternant is (14.7_2), where, by the last formula of (14.8), (VII) = 0. After adjoining (14.7_2), we form the alternant (14.9). Here, because of (I) = 0, (VII) = 0, the coefficients of σ_z, $\sigma_{z'}$ are zero, while also (VIII) = 0 by (14.10). Hence (14.9) reduces to

$$(\text{IX}')\sigma = 0, \tag{15.2}$$

where (IX′) denotes (IX) without its last term, this vanishing because of (I) = 0.

Accordingly, we are led to subdivide Case IIb1 into

$$
\begin{aligned}
&\text{Case IIb1}'. \quad (\text{I}) = 0, \quad (\text{IX}') = 0; \\
&\text{Case IIb1}''. \quad (\text{I}) = 0, \quad (\text{IX}') \neq 0.
\end{aligned} \tag{15.3}
$$

In Case IIb1′, the equations (14.4_2), (14.4_4), (14.7_2) form a complete linear differential system of first order in σ alone—the passivity condition (14.9), or (15.2), disappearing identically. If any solution σ of this system is substituted in the remaining equations of (14.4, 7), these form a complete system for ρ. Hence a solution of the total system in ρ, σ exists, of the generality $\infty^{2f(2)}$. This result is expressed in Theorem IV of §3.

In Case IIb1″, the equation (15.2) implies $\sigma = 0$, in contradiction with the prescribed inequation $\sigma \neq 0$. Hence, in this case the given family of curves is nonextremal.

Example of Case IIb1′.

$$y'' = z, \qquad z'' = 0. \tag{15.4}$$

Here by (7.6), (7.9),

$$
\begin{aligned}
A &= -2, \quad B = 0, \quad C = 0; \\
A_1 &= 0, \quad B_1 = 0, \quad C_1 = 0.
\end{aligned} \tag{15.5}
$$

Hence (9.1) is verified, and we are in Case II.

The quadratic (9.6) is here $\lambda^2 = 0$, with the double root $\lambda = 0$. Since this obeys (I) = 0, we have all the conditions of Case IIb1.

The differential system (14.4) is here

$$\frac{d\rho}{dx} = 0, \qquad \frac{d\sigma}{dx} = 0, \tag{15.6}$$
$$\rho_{y'} = \sigma_{z'}, \qquad \sigma_{y'} = 0.$$

To this we adjoin the passivity conditions corresponding to (14.7):

$$\rho_y = \sigma_z, \qquad \sigma_y = 0. \tag{15.7}$$

The unique passivity condition of this system, corresponding to (14.9), is found to disappear identically, since $(IX')=0$; thus (15.6, 7) form a completely integrable system.

It is profitable here, as in many other cases, to use the finite equations of the extremals as well as their differential equations (15.4). Integrating the latter, we find

$$\begin{aligned} y &= \tfrac{1}{6}a_1x^3 + \tfrac{1}{2}a_2x^2 + a_3x + a_4, & z &= a_1x + a_2; \\ y' &= \tfrac{1}{2}a_1x^2 + a_2x + a_3, & z' &= a_1. \end{aligned} \tag{15.8}$$

These equations may be solved for the four arbitrary constants a_1, a_2, a_3, a_4:

$$\begin{aligned} a_1 &= z', \\ a_2 &= z - z'x, \\ a_3 &= y' - xz + \tfrac{1}{2}z'x^2, \\ a_4 &= y - y'x + \tfrac{1}{2}zx^2 - \tfrac{1}{6}z'x^3. \end{aligned} \tag{15.9}$$

The system (15.6, 7) may be transformed from x, y, z, y', z' as independent variables to x, a_1, a_2, a_3, a_4; it becomes

$$\begin{aligned} \frac{\partial\rho}{\partial x} &= 0, \qquad \frac{\partial\sigma}{\partial x} = 0, \\ \rho_3 - x\rho_4 &= \sigma_1 - x\sigma_2 + \tfrac{1}{2}x^2\sigma_3 - \tfrac{1}{6}x^3\sigma_4, \\ \rho_4 &= \sigma_2 - x\sigma_3 + \tfrac{1}{2}x^2\sigma_4, \\ \sigma_3 - x\sigma_4 &= 0, \\ \sigma_4 &= 0. \end{aligned} \tag{15.10}$$

Here the symbol for partial differentiation, $\partial/\partial x$, is used instead of d/dx, since now, in differentiation as to x, this variable alone changes in value while the other independent variables a_1, a_2, a_3, a_4 are held fixed. The subscripts 1, 2, 3, 4 denote differentiation as to the latter variables respectively.

By the first two equations of (15.10), $\rho=\rho(a_1, a_2, a_3, a_4)$, $\sigma=\sigma(a_1, a_2, a_3, a_4)$. The remaining four equations can be simplified as follows, with the disappearance of x:

$$\rho_3 = \sigma_1, \qquad \rho_4 = \sigma_2, \tag{15.11}$$
$$\sigma_3 = 0, \qquad \sigma_4 = 0.$$

By the last two,

$$\sigma = f(a_1, a_2), \tag{15.12}$$

f being an arbitrary function; and then by the first two:

$$\rho = a_3 f_1(a_1, a_2) + a_4 f_2(a_1, a_2) + g(a_1, a_2), \tag{15.13}$$

where $f_1 \equiv \partial f/\partial a_1$, $f_2 \equiv \partial f/\partial a_2$, and g is a second arbitrary function.

By (14.2), since $\lambda = 0$, we have $L = 0$, $M = \sigma$, $N = \rho$; hence by (15.12, 13) and the substitutions (15.9), we have L, M, N as functions of x, y, z, y', z'. The integrand ϕ can then be found by the method of §5.

Example of Case IIb1''. Let $G = 0$; then $C = 0$. Let F be such that

$$B = 0, \qquad A \neq 0, \qquad F_{y'y'y'} \neq 0. \tag{15.14}$$

These conditions are consistent, since the partial differential equation $B = 0$ obeyed by F, that is,

$$F_{y'y'} = \left\{2F_y + \tfrac{1}{2}F_{y'}^2 - F_{xy'} - y'F_{y'y'} - z'F_{zy'} - GF_{y'z'}\right\} \div F,$$

is of the Cauchy normal form, and therefore has a solution corresponding to arbitrarily given analytic functions f, g if we prescribe that

$$f(x, y, z, z') = F(x, y, z, 0, z'), \qquad g(x, y, z, z') = F_{y'}(x, y, z, 0, z').$$

It is evidently possible to choose these two arbitrary functions so that the inequations in (15.14) are satisfied.

By (7.9) and our stated conditions, we have

$$B_1 = 0, \qquad C_1 = 0;$$

therefore, since also $B = 0$, $C = 0$, $A \neq 0$, we are in Case II. The fundamental quadratic, $A\xi^2 + B\xi + C = 0$, has the double root $\lambda = 0$, for which $\lambda\lambda_{z'} - \lambda_{y'} = 0$; we thus have the conditions of Case IIb1.

Further, with reference to (14.8), we have

$$\text{(I)} = 0, \qquad \text{(II)} = 0, \qquad \text{(VI)} = -\tfrac{1}{2}F_{y'y'y'}, \qquad \text{(VII)} = 0;$$

therefore (14.4$_4$), (14.7$_2$) become

$$\sigma_{y'} = 0, \qquad \sigma_y = -\tfrac{1}{2}F_{y'y'y'}\sigma. \tag{15.15}$$

Taking $\partial/\partial y'$ of the second equation and using the first, we get as the analogue of (14.9):

$$F_{y'y'y'y'}\sigma = 0; \tag{15.16}$$

consequently, by the stipulation $F_{y'y'y'} \neq 0$ of (15.14), we deduce $\sigma = 0$, in contradiction with the prescribed inequation $\sigma \neq 0$.

16. **Case IIb2.** In this case, where $(\mathrm{I}) \neq 0$, the passivity condition (14.9) of the system (14.4, 7) is not verified identically, nor does it offer any direct contradiction with the prescribed inequation $\sigma \neq 0$; therefore we adjoin (14.9) and proceed to form further passivity conditions.

With the abbreviation

$$(\mathrm{X}) = 2(\mathrm{I})(\mathrm{IV}) - (\mathrm{VII}), \tag{16.1}$$

we rewrite (14.9) as

$$\left[2(\mathrm{I})\frac{\partial}{\partial z} + (\mathrm{X})\frac{\partial}{\partial z'}\right]\sigma = -(\mathrm{VIII})\rho - (\mathrm{IX})\sigma. \tag{16.2}$$

Similarly, we write (14.4_4) in the operational form

$$\left[\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}\right]\sigma = (\mathrm{I})\rho + 2\lambda_{z'}\sigma. \tag{16.3}$$

Forming then the alternant of (16.2), (16.3), and using the other equations of our system, we obtain a new first order equation of the form

$$(\mathrm{XI})\sigma_z + (\mathrm{XII})\sigma_{z'} = (\mathrm{I})\left[2(\mathrm{I})\rho_z + (\mathrm{X})\rho_{z'}\right] + (\mathrm{XIII})\rho + (\mathrm{XIV})\sigma, \tag{16.4}$$

where the coefficients represented by Roman numerals can be expressed by explicit formulas in terms of the preceding Roman numerals and their partial derivatives, namely:

$$\begin{aligned}
(\mathrm{XI}) &= -2\left[(\mathrm{I})_{y'} - \lambda(\mathrm{I})_{z'}\right], \\
(\mathrm{XII}) &= -2(\mathrm{I})\lambda_z - (\mathrm{X})\lambda_{z'} - (\mathrm{X})_{y'} + \lambda(\mathrm{X})_{z'} - (\mathrm{VIII}), \\
(\mathrm{XIII}) &= 2(\mathrm{I})(\mathrm{I})_z + (\mathrm{X})(\mathrm{I})_{z'} + (\mathrm{VIII})_{y'} - \lambda(\mathrm{VIII})_{z'} - \lambda_{z'}(\mathrm{VIII}) \\
&\quad + (\mathrm{I})(\mathrm{IX}), \\
(\mathrm{XIV}) &= 2\left[2(\mathrm{I})\lambda_{zz'} + (\mathrm{X})\lambda_{z'z'}\right] + (\mathrm{IX})_{y'} - \lambda(\mathrm{IX})_{z'}.
\end{aligned} \tag{16.5}$$

Similarly, we may write (14.7_3) in the operational form

$$\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (\mathrm{II})\frac{\partial}{\partial z'}\right]\sigma = (\mathrm{VI})\sigma + (\mathrm{VII})\rho, \tag{16.6}$$

and build the alternant of this with (16.2); we get, after reduction with the use of the other equations of the system:

$$(\mathrm{XV})\sigma_z + (\mathrm{XVI})\sigma_{z'} = (\mathrm{VII})\left[2(\mathrm{I})\rho_z + (\mathrm{X})\rho_{z'}\right] + (\mathrm{XVII})\rho + (\mathrm{XVIII})\sigma, \tag{16.7}$$

where

$$(XV) = -\left[2(I)\frac{\partial}{\partial z} + (X)\frac{\partial}{\partial z'}\right]\lambda$$
$$- 2\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (II)\frac{\partial}{\partial z'}\right](I) - (VIII),$$
$$(XVI) = -\left[2(I)\frac{\partial}{\partial z} + (X)\frac{\partial}{\partial z'}\right](II)$$
$$-\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (II)\frac{\partial}{\partial z'}\right](X) - (IV)(VIII),$$

(16.8)
$$(XVII) = \left[2(I)\frac{\partial}{\partial z} + (X)\frac{\partial}{\partial z'}\right](VII)$$
$$+\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (II)\frac{\partial}{\partial z'}\right](X)$$
$$+ (III)(VIII) - (VI)(VIII) + (VII)(IX),$$
$$(XVIII) = \left[2(I)\frac{\partial}{\partial z} + (X)\frac{\partial}{\partial z'}\right](VI)$$
$$+\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (II)\frac{\partial}{\partial z'}\right](IX) + (V)(VIII).$$

The passivity condition formed by taking the alternant of (14.9) with (14.4_3) is found to be disappear identically in virtue of the other equations of the system. Hence (16.4), (16.7) represent all the integrability conditions of the differential system (14.4, 7, 9); and clearly this system is not passive, since (16.4) certainly does not disappear identically, the coefficient of ρ_z in this equation being $(I)^2$, which is $\neq 0$ by the hypothesis of Case IIb2.

Forming next the linear combination: (VII) times (16.4) minus (I) times (16.7), we eliminate the derivatives of ρ and get

(16.9) $$(XIX)\sigma_z + (XX)\sigma_{z'} = (XXI)\rho + (XXII)\sigma,$$

where

(16.10) $$(XIX) = (VII)(XI) - (I)(XV),$$

with similar expressions for (XX), (XXI), (XXII).

We now compare (16.9) with (16.2). First, suppose that these are independent as linear algebraic equations in σ_z, $\sigma_{z'}$, ρ, σ, that is, that not all the two-rowed determinants in the following matrix vanish:

(16.11) $$\left\| \begin{matrix} 2(I) & (X) & -(VIII) & -(IX) \\ (XIX) & (XX) & (XXI) & (XXII) \end{matrix} \right\|.$$

Then we consider two possibilities, namely, the determinant of the first two columns:

$$2(\mathrm{I})(\mathrm{XX}) - (\mathrm{X})(\mathrm{XIX}), \tag{16.12}$$

either does or does not vanish.

If it vanishes, we infer, by linear combination of (16.9) and (16.2), a relation of the form

$$\cdot\rho + \cdot\sigma = 0, \tag{16.13}$$

where the dots represent known coefficients not both zero (for (16.9) and (16.2) are, by hypothesis, linearly independent).

If the coefficient of ρ in (16.13) is zero, the coefficient of σ cannot be zero; it follows that $\sigma=0$, contrary to the prescribed inequation $\sigma\neq 0$. There is then no solution of our differential system.

On the other hand, if the coefficient of ρ in (16.13) is $\neq 0$, we can solve this equation for ρ in the form $\rho = \cdot\sigma$, and substituting in our original system (14.4, 7), we have the classical case of a linear differential system of first order in σ alone—easily settled here by reference to the exactness or non-exactness of a certain readily constructed differential, $\cdot dx + \cdot dy + \cdot dz + \cdot dy' + \cdot dz'$.

Suppose now that the determinant (16.12) does not vanish. Then (16.2), (16.9) are solvable in the form

$$\sigma_z = \cdot\rho + \cdot\sigma, \qquad \sigma_{z'} = \cdot\rho + \cdot\sigma. \tag{16.14}$$

Substituting this in the other equations of our system, (14.4), (14.7), (14.9), (16.4), we find a differential system in ρ, σ of the form:

$$\begin{aligned}
&\frac{d\rho}{dx} = \cdot\rho + \cdot\sigma, && \frac{d\sigma}{dx} = \cdot\sigma,\\
&\rho_y = \cdot\rho_{z'} + \cdot\rho + \cdot\sigma, && \sigma_y = \cdot\rho + \cdot\sigma,\\
&\rho_z = \cdot\rho_{z'} + \cdot\rho + \cdot\sigma, && \sigma_z = \cdot\rho + \cdot\sigma,\\
&\rho_{y'} = \cdot\rho_{z'} + \cdot\rho + \cdot\sigma; && \sigma_{y'} = \cdot\rho + \cdot\sigma,\\
& && \sigma_{z'} = \cdot\rho + \cdot\sigma.
\end{aligned} \tag{16.15}$$

All the passivity conditions which occur in the discussion of this system, as arrived at by formation of alternants of first order differential expressions, are evidently of the form

$$\cdot\rho_{z'} + \cdot\rho + \cdot\sigma = 0.$$

Since there are at most three linearly independent equations of this type, we must arrive, after just a few additional steps, at a definite decision as to the solvability of our differential system and the generality of its solution.

With this, we have disposed of the case of rank 2 of the matrix (16.11). Suppose, finally, that this matrix is of rank 1 (it cannot be of rank 0, since $(\mathrm{I}) \neq 0$).

Then (16.7) is linearly dependent on (16.2) and (16.4), so that we have to adjoin only (16.4) as the sole passivity condition of the previously existent system. We now have in (14.4, 7, 9) and (16.4) a differential system which can be expressed in the following operational form:

$$
\begin{aligned}
&\frac{d\rho}{dx} = \cdot\rho + \cdot\sigma,\\
&\left[\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}\right]\rho = \sigma_{x'} + \cdot\rho\,(^{31}),\\
&\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (\mathrm{II})\frac{\partial}{\partial z'}\right]\rho = \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma,\\
&\left[2(\mathrm{I})\frac{\partial}{\partial z} + (\mathrm{X})\frac{\partial}{\partial z'}\right]\rho = \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma;
\end{aligned}
\tag{16.16}
$$

$$
\begin{aligned}
&\frac{d\sigma}{dx} = \cdot\sigma,\\
&\left[\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}\right]\sigma = \cdot\rho + \cdot\sigma,\\
&\left[\frac{\partial}{\partial y} - \lambda\frac{\partial}{\partial z} - (\mathrm{II})\frac{\partial}{\partial z'}\right]\sigma = \cdot\rho + \cdot\sigma,\\
&\left[2(\mathrm{I})\frac{\partial}{\partial z} + (\mathrm{X})\frac{\partial}{\partial z'}\right]\sigma = \cdot\rho + \cdot\sigma.
\end{aligned}
\tag{16.17}
$$

The important point to observe is that *the operators on ρ and on σ are the same.*

The fourth equation of (16.16) is the new one of this system, and the new passivity conditions are found by forming the alternant of this equation with the other ones of (16.16). The alternant with the first equation of (16.16) turns out to disappear identically in virtue of the other equations of the system.

The remaining alternants are of the first order in the partial derivatives of σ as well as ρ, precisely because of the identity of corresponding operators in (16.16), (16.17). For instance, after forming the alternant of the second and fourth equations in (16.16), we have a term of the form

---

(31) Note that the coefficient of $\sigma_{x'}$ in this equation is unity. The dots in the other equations represent, as usual, calculable known coefficients.

$$(16.18) \qquad \left[\frac{\partial}{\partial y'} - \lambda \frac{\partial}{\partial z'}\right]\frac{\partial}{\partial z'}\sigma,$$

apparently of the second order; but this is equal to

$$\left\{\frac{\partial}{\partial y'} - \lambda \frac{\partial}{\partial z'}, \frac{\partial}{\partial z'}\right\}\sigma + \frac{\partial}{\partial z'}\left[\frac{\partial}{\partial y'} - \lambda \frac{\partial}{\partial z'}\right]\sigma,$$

where the braces denote the alternant of the two operators separated by the comma—and this alternant is of the first order. The second term is, by the second equation of (16.17), equal to

$$\frac{\partial}{\partial z'}(\cdot\rho + \cdot\sigma),$$

a differential expression of first order. Hence the term (16.18) is equal to a first order expression in the partial derivatives of ρ, σ, and a similar reduction applies to other apparently second order terms arising in the formation of our passivity conditions.

In summary, we obtain two passivity conditions of the form

$$(16.19) \qquad \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma = 0, \qquad \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma = 0.$$

If these two are linearly independent, then either they can be solved for $\rho_{x'}, \sigma_{x'}$:

$$(16.20) \qquad \rho_{x'} = \cdot\rho + \cdot\sigma, \qquad \sigma_{x'} = \cdot\rho + \cdot\sigma,$$

or else a non-identically vanishing relation of the form $\cdot\rho + \cdot\sigma = 0$ can be deduced from them by linear combination.

In the latter case, the discussion of our system is easily finished along lines already discussed (cf. (16.13)). In the former case, if we substitute (16.20) in (16.16), (16.17), we obtain a differential system in which *every* first partial derivative of ρ, σ is expressed in the form $\cdot\rho + \cdot\sigma$. Then all the passivity conditions of this system are clearly of the form $\cdot\rho + \cdot\sigma = 0$, so that the discussion of the system is easily concluded along the lines of standard theory[32].

If both equations (16.19) disappear identically, the system (16.16, 17) is completely integrable.

The only remaining case is where the matrix of the coefficients in (16.19) is of rank 1, so that these equations amount to a single linearly independent one:

$$(16.21) \qquad \cdot\rho_{x'} + \cdot\sigma_{x'} + \cdot\rho + \cdot\sigma = 0.$$

If the coefficient of $\rho_{x'}$ is zero, then either we have $\cdot\rho + \cdot\sigma = 0$, a case already

[32] Cf., as in connection with (13.15), Eisenhart, loc. cit.

disposed of, or else $\sigma_{z'} = \cdot\rho + \cdot\sigma$. In the latter case, we obtain by substitution in (16.16, 17) a system of the form (16.15), already discussed.

Accordingly, we may suppose (16.21) solved for $\rho_{z'}$:

$$\rho_{z'} = \cdot\sigma_{z'} + \cdot\rho + \cdot\sigma. \tag{16.22}$$

Let us form the alternant of this with the second equation of (16.16):

$$\frac{\partial}{\partial z'}(16.16_2) - \left[\frac{\partial}{\partial y'} - \lambda\frac{\partial}{\partial z'}\right](16.22). \tag{16.23}$$

Then we get a passivity condition of the form

$$\sigma_{z'z'} = \cdot\sigma_{z'} + \cdot\rho + \cdot\sigma, \tag{16.24}$$

solvable for $\sigma_{z'z'}$, whose coefficient in (16.23) is evidently 1. There are two additional passivity conditions, obtained by building the alternant of (16.22) with the third and fourth equations of (16.16)([33]); these are of the form

$$\cdot\sigma_{z'z'} + \cdot\sigma_{z'} + \cdot\rho + \cdot\sigma = 0.$$

Substituting in this the expression (16.24) for $\sigma_{z'z'}$, we get two equations of the type

$$\cdot\sigma_{z'} + \cdot\rho + \cdot\sigma = 0, \qquad \cdot\sigma_{z'} + \cdot\rho + \cdot\sigma = 0. \tag{16.25}$$

Unless both of these disappear identically, we can deduce either $\cdot\rho + \cdot\sigma = 0$, $\sigma_{z'} = \cdot\rho + \cdot\sigma$. In either case we find ourselves in situations already discussed([34]), and our investigation finishes along simple lines of standard theory.

If, however, both equations (16.25) disappear identically, we have to adjoin (16.22), (16.24) as the passivity conditions of the system (16.16, 17). Then, clearly, the extended system has only the three parametric derivatives $\sigma_{z'}$, ρ, σ; consequently all further passivity conditions must be of the form

$$\cdot\sigma_{z'} + \cdot\rho + \cdot\sigma = 0.$$

Evidently, our investigation now comes to its conclusion by reduction to cases already considered.

We may conclude by observing that throughout the discussion of Case IIb, even in the most unfavorable eventualities, the only derivative of an order higher than the first which can occur is $\sigma_{z'z'}$.

17. **Case III.** This is the case of rank *two* of the matrix Δ, that is:

$$\begin{vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} = 0, \qquad \left\| \begin{matrix} A & B & C \\ A_1 & B_1 & C_1 \end{matrix} \right\| \neq 0. \tag{17.1}$$

([33]) The condition obtained by alternating with d/dx (first equation of (16.16)) disappears in virtue of the other equations in our system.

([34]) As is readily seen in the latter case by combining (16.26), (16.22) with (16.16, 17).

Let us denote the determinants in the latter matrix as follows:

$$\Delta_1 \equiv BC_1 - CB_1, \quad \Delta_2 \equiv CA_1 - AC_1, \quad \Delta_3 \equiv AB_1 - BA_1; \tag{17.2}$$

then according to (17.1), at least one of $\Delta_1, \Delta_2, \Delta_3$ is not zero.

It follows, by the linear equations $AL+BM+CN=0$, $A_1L+B_1M+C_1N=0$, of (7.5), (7.7), that a proportionality factor ρ, function of x, y, z, y', z', exists such that

$$L = \rho\Delta_1, \qquad M = \rho\Delta_2, \qquad N = \rho\Delta_3. \tag{17.3}$$

ρ is now the sole unknown function of our problem.

By (17.3),

$$LN - M^2 = \rho^2(\Delta_1\Delta_3 - \Delta_2^2); \tag{17.4}$$

hence, by the fundamental inequation $LN-M^2\neq 0$ of (7.5), we must have

$$D \equiv \Delta_1\Delta_3 - \Delta_2^2 \neq 0 \tag{17.5}$$

—otherwise, we can conclude immediately that the given curve family is non-extremal.

We are thus led to make the following subdivision of Case III:

(17.6) Case IIIa. $D \neq 0$,

(17.7) Case IIIb. $D = 0$([35]).

Example of Case IIIb.

$$y'' = y^2 + z^2, \qquad z'' = 0. \tag{17.8}$$

We calculate by (7.6), (7.9):

$$\begin{aligned} &A = -4z, && B = 4y, && C = 0; \\ &A_1 = -4z', && B_1 = 4y', && C_1 = 0; \\ &A_2 = 0, && B_2 = 4(y^2+z^2), && C_2 = 0. \end{aligned} \tag{17.9}$$

Here

$$\Delta_1 = 0, \qquad \Delta_2 = 0, \qquad \Delta_3 = 16(yz' - zy');$$

therefore $D=0$.

The curve family (17.8) is nonextremal.

Case IIIa. Let us substitute (17.3) for L, M, N in the differential system

([35]) Geometrically expressed, (17.3), where ρ is regarded as a parameter, represents a line in LMN-space passing through the vertex of the critical cone. The distinction between Cases IIIb, and IIIa, respectively, consists in whether this line is or is not an element of the critical cone.

(7.5) as well as in the equations (8.8), (8.9) of its extended system. Then we obtain exactly five independent linear equations in the five first order derivatives of ρ; and these equations can be solved for the derivatives in question, giving

$$\rho_x = E_1\rho, \quad \rho_y = E_2\rho, \quad \rho_z = E_3\rho, \quad \rho_{y'} = E_4\rho, \quad \rho_{z'} = E_5\rho. \tag{17.10}$$

Here the coefficients E_i are calculable known functions, being rational expressions in the partial derivatives of F, G each having D for denominator; accordingly, the E_i exist as determinate quantities in virtue of the stipulation $D \neq 0$.

By (17.4), (17.5), the fundamental inequation $LN - M^2 \neq 0$ is seen to be expressed precisely by requiring a solution ρ of the system (17.10) other than the trivial one, zero:

$$\rho \not\equiv 0. \tag{17.11}$$

For the existence of such a ρ, the exactness of the differential

$$E_1dx + E_2dy + E_3dz + E_4dy' + E_5dz' \tag{17.12}$$

is evidently a necessary and sufficient condition. If, in case of exactness, the integral of this differential is $\zeta(x, y, z, y', z')$, then

$$\rho = e^{\zeta}. \tag{17.13}$$

ζ is determined up to an additive constant, and ρ, therefore, up to a constant factor.

After ρ has been found, we substitute in (17.3) to obtain L, M, N, and therefrom we find ϕ by the method of §5. L, M, N will be determined up to a constant factor, since that is the case with ρ. By the relations (7.3) and by §5, the integrand ϕ is therefore determined up to a constant factor, besides the usual additive total derivative as to x—thus, according to the convention associated with (3.7), ϕ is essentially uniquely determined.

The results of this section are expressed in Theorems VI and VII of §3.

Example of Case IIIa.

$$y'' = z^2, \qquad z'' = y^2. \tag{17.14}$$

By calculation, using (7.6), (7.9):

$$\begin{aligned} &A = -4z, && B = 0, && C = 4y; \\ &A_1 = -4z', && B_1 = 0, && C_1 = 4y'; \\ &A_2 = -4y^2, && B_2 = 0, && C_2 = 4z^2. \end{aligned} \tag{17.15}$$

The conditions (17.1) and (17.5) are immediately verifiable.

Since $\Delta_1 = 0$, $\Delta_3 = 0$, equations (17.3) give $L = 0$, $N = 0$.

By substitution in the differential system (7.5), (8.8), (8.9), this becomes

the following system for M (which, in this example, plays the part of ρ):

$$M_z = 0, \quad M_y = 0, \quad M_z = 0, \quad M_{y'} = 0, \quad M_{z'} = 0. \tag{17.16}$$

Therefore $M=$const., and since a constant factor is inessential, we may say $M=1$. Knowing now L, M, N, we can easily find ϕ by the method of §5; the result is

$$\phi = y'z' + \tfrac{1}{2}y^2 + \tfrac{1}{2}z^2, \tag{17.17}$$

and this is essentially unique, that is, within the slight possibility of modification expressed by (3.7).

18. **Case IV.** This is the case where the determinant of the fundamental matrix Δ does not vanish:

$$\begin{vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} \neq 0, \tag{18.1}$$

that is, Δ is of rank three.

Since (18.1) contradicts the necessary condition (7.10) for an extremal family, any curve family coming under Case IV is certainly nonextremal.

Example.

$$-y'' = y^2 + z^2, \quad z'' = y. \tag{18.2}$$

By calculation, with reference to (7.6), (7.9):

$$\begin{aligned} A &= -4z, & B &= 4y, & C &= 2; \\ A_1 &= -4z', & B_1 &= 4y', & C_1 &= 0; \\ A_2 &= -4y, & B_2 &= 4(y^2+z^2), & C_2 &= 0. \end{aligned} \tag{18.3}$$

The determinant of the matrix Δ is not equal to zero. The curve family (18.2) is therefore non-extremal.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

SUR L'EXISTENCE DE FONCTIONS ENTIÈRES SATISFAISANT À CERTAINES CONDITIONS LINÉAIRES

PAR

G. PÓLYA

1. Le problème. Nous donnons deux suites infinies de nombres complexes

$$a_0, a_1, a_2, \cdots, a_n, \cdots,$$

$$A_0, A_1, A_2, \cdots, A_n, \cdots$$

et une suite infinie de nombres entiers non-négatifs

$$\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_n, \cdots.$$

Nous cherchons une fonction entière $F(z)$ qui satisfait aux conditions, en nombre infini,

$$F^{(\alpha_0)}(a_0) = A_0, F^{(\alpha_1)}(a_1) = A_1, \cdots, F^{(\alpha_n)}(a_n) = A_n, \cdots; \tag{1}$$

on suppose naturellement que la même dérivée de $F(z)$ au même point ne figure qu'une fois dans les conditions (1), supposition qu'on peut mettre sous la forme:

$$\text{si } a_k = a_l \text{ et } k < l, \text{ alors } \alpha_k < \alpha_l. \tag{2}$$

Existe-t-il une fonction entière $F(z)$ satisfaisant aux conditions (1)?

La fonction entière $F(z)$ existe certainement si

$$\lim_{n\to\infty} a_n = \infty$$

et les nombres α_n et A_n sont, excepté (2), arbitraires. C'est un résultat classique(1).

Très différent du cas classique mentionné est celui où la suite des a_n est *bornée*, tandis que

Presented to the Society September 12, 1940; received by the editors July 1, 1940.

(1) Voir *Guichard*, Annales Scientifique de l'École Normale Supérieure, (3), Tome 1 (1884), pp. 427–432; énoncé particulier, démonstration (valable dans le cas général!) basée sur les théorèmes classiques de Weierstrass et de Mittag-Leffler. Voir aussi *G. Pólya*, Commentarii Mathematici Helvetici, Tome 11 (1938), pp. 234–252, surtout "Aufgabe 3," p. 245 et "Aufgabe 5," p. 249; énoncé général, démonstration faisant usage d'un système infini d'équations linéaires. (Je saisis l'occasion pour donner deux rectifications à ce travail: 1° En l'écrivant, je n'ai pas connu deux travaux de *M. Eidelheit*, Studia Mathematica, Tome 6 (1936), pp. 139–148 et Tome 7 (1937), pp. 150–154. 2° L'énoncé du Théorème II, p. 241, doit être rectifié: au lieu de la condition B'' (p. 241) on doit supposer la condition plus restreinte, mentionnée sous 3 (p. 241).

$$\lim_{n\to\infty} \alpha_n = \infty. \tag{3}$$

Pour que la fonction entière $F(z)$ existe dans ce cas-là, il est nécessaire que la suite des A_n satisfasse à la condition

$$\lim_{n\to\infty} \left| \frac{A_n}{\alpha_n!} \right|^{1/\alpha_n} = 0; \tag{4}$$

la nécessité de (4) est une conséquence immédiate d'une inégalité classique de Cauchy. Mais la condition (4), jointe aux autres mentionnées, est encore loin d'être suffisante pour l'existence de $F(z)$ (voir §6).

Je ne peux démontrer l'existence de $F(z)$ que dans un cas beaucoup plus restreint que le précédent: les a_n ne sont pas seulement bornés, mais ils sont supposés n'être capables que d'un *nombre fini de valeurs* $c_1, c_2, \cdots, c_q$; je suppose, en outre, une certaine périodicité de période p. Voici l'énoncé complet.

THÉORÈME. *On donne*

un nombre entier p;

q points différents dans le plan complexe $c_1, c_2, \cdots, c_q$;

la suite infinie $a_0, a_1, a_2, \cdots$ *dont chaque terme est égal à un certain terme de la suite finie* $c_1, c_2, \cdots, c_q$—*on suppose que*

$$a_{n+p} = a_n \qquad \text{pour } n = 0, 1, 2, \cdots; \tag{5}$$

la suite infinie d'entiers $\alpha_0, \alpha_1, \alpha_2, \cdots$—*on suppose* (2) *et*

$$\alpha_{n+p} = \alpha_n + p \qquad \text{pour } n = 0, 1, 2, \cdots, \tag{6}$$

$$0 \leqq \alpha_0 \leqq \alpha_1 \leqq \alpha_2 \leqq \cdots. \tag{7}$$

On suppose en outre $a_0, a_1, a_2, \cdots, a_{p-1}, \alpha_0, \alpha_1, \cdots, \alpha_{p-1}$ *choisis de telle façon qu'il n'existe aucun polynôme* $P(z)$ *de degré inférieur à* p *et non identiquement nul qui satisfasse aux* p *équations*

$$P^{(\alpha_0)}(a_0) = 0, P^{(\alpha_1)}(a_1) = 0, \cdots, P^{(\alpha_{p-1})}(a_{p-1}) = 0. \tag{8}$$

On donne enfin la suite infinie $A_0, A_1, A_2, \cdots$; *on suppose* (4).

Sous ces conditions, il existe une fonction entière $F(z)$ *qui satisfait aux conditions* (1).

Les conditions de ce théorème sont un peu complexes; nous les discuterons au §2, et, après certaines préparations aux §§3 et 4, nous démontrerons le théorème au §5.

Notre théorème généralise deux cas particuliers traités par J. M. Whittaker; la démonstration du théorème que je donnerai résulte de l'analyse attentive des démonstrations données par Whittaker et le résultat du théorème

constitue une réponse partielle à une question posée par Whittaker(2). Certaines questions voisines ont été traitées par Poritsky(3) et Gontcharoff(4).

2. **Discussion des conditions.** En vertu des conditions (5) et (6), les équations (1) prennent la forme suivante:

$$
\begin{aligned}
&F^{(\alpha_0)}(a_0) = A_0, \quad F^{(\alpha_1)}(a_1) = A_1, \quad \cdots, F^{(\alpha_{p-1})}(a_{p-1}) = A_{p-1},\\
&F^{(\alpha_0+p)}(a_0) = A_p, \quad F^{(\alpha_1+p)}(a_1) = A_{1+p}, \cdots, F^{(\alpha_{p-1}+p)}(a_{p-1}) = A_{2p-1},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,\\
&F^{(\alpha_0+mp)}(a_0) = A_{mp}, \cdots,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots.
\end{aligned}
\tag{1'}
$$

Observons encore que (2) et (7) sont évidemment compatibles et que (6) implique (3).

Quelle est la signification de la condition relative aux équations (8)? $P(z)$ est un polynôme de degré $<p$, donc d'un degré $\leqq p-1$; il peut être mis sous la forme

$$P(z) = u_0 + u_1 z + u_2 z^2 + \cdots + u_{p-1} z^{p-1}.$$

(8) constitute un système de p équations, linéaires et homogènes, pour déterminer les p inconnues $u_0, u_1, \cdots, u_{p-1}$. Notre condition demande que le polynôme identiquement nul soit l'unique solution des équations (8), donc que

$$u_0 = u_1 = u_2 = \cdots = u_{p-1} = 0$$

soit l'unique solution du système linéaire et homogène mentionné. Notre condition consiste donc finalement en ceci: *le déterminant* Δ *du système est* $\neq 0$.

Toutes les lignes du déterminant Δ sont de la forme

$$0, \cdots, 0, \alpha!, \frac{(\alpha+1)!}{1!}a, \frac{(\alpha+2)!}{2!}a^2, \cdots, \frac{(p-1)!}{(p-\alpha-1)!}a^{p-\alpha-1}. \tag{9}$$

Le nombre des 0 au commencement de la ligne est α; pour obtenir les lignes successives de Δ, il faut poser $\alpha = \alpha_k$, $a = a_k$ et, après cela, $k = 0, 1, 2, \cdots, p-1$.

En particulier, au commencement des lignes successives de Δ, nous trouvons $\alpha_0, \alpha_1, \cdots, \alpha_{p-1}$ éléments évanouissants; ces nombres vont en croissant, d'après (7). Puisque $\Delta \neq 0$, nous avons nécessairement

(2) *J. M. Whittaker*, Proceedings of the London Mathematical Society, (2), Tome 36 (1934), pp. 451–469. Voir le Théorème 3 (p. 457), son analogue (p. 458) et le théorème hypothétique formulé au début de la page 465. Voir aussi *J. M. Whittaker, Interpolatory Function Theory*, Cambridge, 1935, p. 51.

(3) *H. Poritsky*, ces Transactions, Tome 34 (1923), pp. 274–331.

(4) *W. Gontcharoff*, Communications de la Société Mathématique de Kharkoff, (4), Tome 5 (1932), pp. 67–85.

$$\alpha_0 = 0, \alpha_1 \leqq 1, \alpha_2 \leqq 2, \cdots, \alpha_{p-1} \leqq p-1. \tag{10}$$

Occupons-nous encore brièvement des cas particuliers les plus simples. Si $p=1$, le tableau (1′) se réduit à une seule colonne et, puisque $\alpha_0=0$ en vertu de (10), le problème se réduit à la construction d'une fonction entière dont toutes les dérivées sont données en un point donné—problème qui est sans difficulté. Lorsque

$$p = 2, \qquad a_0 = c_1 \neq a_1 = c_2,$$

le tableau (1′) consiste en 2 colonnes et il n'y a, en vertu de (10), que deux possibilités:

$$\alpha_0 = 0, \ \alpha_1 = 0 \quad \text{ou} \quad \alpha_0 = 0, \ \alpha_1 = 1.$$

Dans le premier de ces cas, toutes les dérivées d'ordre pair sont données en deux points différents; dans le second cas, on donne dans un point toutes les dérivées d'ordre pair et dans un second point, différent du premier, toutes les dérivées d'ordre impair. Voilà les deux cas dans lesquels Whittaker a démontré l'existence de la fonction entière $F(z)$.

Voici deux cas pour p quelconque, dans lesquels Δ ne s'annulle pas:

(I) $\alpha_0=\alpha_1=\cdots=\alpha_{p-1}=0$, les nombres $a_0, a_1, \cdots, a_{p-1}$ sont distincts, Δ se réduit à un déterminant de Vandermonde;

(II) $\alpha_0=0, \alpha_1=1, \cdots, \alpha_{p-1}=p-1$, les nombres $a_0, a_1, \cdots, a_{p-1}$ sont quelconques, Δ se réduit au produit de ses éléments diagonaux.

Le cas (I) a été envisagé par Poritsky et le cas (II) par Gontcharoff; ces deux auteurs ont eu un but différent, ils ne se sont pas occupés du problème d'existence qui nous intéresse ici, mais d'un certain problème de convergence(5); toutefois quelques-unes des formules qu'ils ont développées peuvent servir de point de départ pour les formules plus générales que nous allons développer.

3. **Accès heuristique.** Les conditions (1) constituent une sorte d'extension à l'infini du problème classique qui est résolu par la formule d'interpolation de Lagrange. Faisons subir à cette formule l'extension analogue; nous arrivons à la formule

$$F(z) = \sum_{n=0}^{\infty} F^{(\alpha_n)}(a_n) P_n(z) \tag{11}$$

où $P_n(z)$ est un certain polynôme dont la nature dépend des deux suites $a_0, a_1, \cdots$ et $\alpha_0, \alpha_1, \cdots$. Appliquons (11) à la fonction entière simple

$$F(z) = e^{zz}$$

(5) Il s'agit de la convergence de la série (11), dans les cas particuliers respectifs; voir loc. cit. (3) et (4). Whittaker s'est aussi occupé de la convergence de cette série dans les deux cas particuliers qu'il a envisagés, loc. cit. (3). L'étude de la convergence de la série (11) générale fera partie d'une Thèse que S. Schmidli prépare à l'Ecole Polytechnique Fédérale de Zürich.

qui dépend du paramètre complexe s; nous arrivons à la formule

$$e^{zs} = \sum_{n=0}^{\infty} s^{\alpha_n} e^{\alpha_n z} P_n(z)$$

que nous écrirons, en tenant compte des conditions (5) et (6), ainsi ($n = pm + k$, passage de (1) à (1')):

$$\begin{aligned} e^{zs} &= \sum_{k=0}^{p-1} s^{\alpha_k} e^{\alpha_k z} \sum_{m=0}^{\infty} s^{pm} P_{pm+k}(z) \\ &= s^{\alpha_0} e^{\alpha_0 z} L_0 + s^{\alpha_1} e^{\alpha_1 z} L_1 + \cdots + s^{\alpha_{p-1}} e^{\alpha_{p-1} z} L_{p-1}. \end{aligned} \tag{12}$$

Nous avons posé pour abréger

$$L_k = L_k(z, s) = \sum_{m=0}^{\infty} s^{pm} P_{pm+k}(z), \tag{13}$$

$k = 0, 1, 2, \cdots, p-1$. Observons qu'en posant

$$e^{2\pi i/p} = \omega \tag{14}$$

nous avons

$$L(z, \omega s) = L(z, s). \tag{15}$$

Par application réitérée de (15) à (12) nous arrivons à un système de p équations linéaires pour les p fonctions $L_0, L_1, \cdots, L_{p-1}$:

$$\begin{aligned} e^{zs} &= s^{\alpha_0} e^{\alpha_0 z} L_0 + \cdots + s^{\alpha_{p-1}} e^{\alpha_{p-1} z} L_{p-1}, \\ e^{z\omega s} &= (\omega s)^{\alpha_0} e^{\alpha_0 \omega z} L_0 + \cdots + (\omega s)^{\alpha_{p-1}} e^{\alpha_{p-1} \omega z} L_{p-1}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \\ e^{z\omega^{p-1} s} &= (\omega^{p-1} s)^{\alpha_0} e^{\alpha_0 \omega^{p-1} z} L_0 + \cdots + (\omega^{p-1} s)^{\alpha_{p-1}} e^{\alpha_{p-1} \omega^{p-1} z} L_{p-1}. \end{aligned} \tag{16}$$

4. Définition et propriétés de quelques fonctions auxiliaires. Effectivement le développement (11) ne joue aucun rôle dans la démonstration qui suit; il nous a donné seulement l'occasion d'introduire les éléments analytiques qui y jouent un rôle. Nous allons définir certaines fonctions auxiliaires qui nous seront utiles, en suivant les suggestions du no. précédent, mais officiellement, nous n'utiliserons aucunement les calculs qui y étaient basés sur le développement (hypothétique!) (11).

I. Nous définissons $D(s)$ par l'équation

$$D(s) = \begin{vmatrix} s^{\alpha_0} e^{\alpha_0 s} & s^{\alpha_1} e^{\alpha_1 s} & \cdots & s^{\alpha_{p-1}} e^{\alpha_{p-1} s} \\ (\omega s)^{\alpha_0} e^{\alpha_0 \omega s} & (\omega s)^{\alpha_1} e^{\alpha_1 \omega s} & \cdots & (\omega s)^{\alpha_{p-1}} e^{\alpha_{p-1} \omega s} \\ \cdot & \cdot & \cdots & \cdot \\ (\omega^{p-1} s)^{\alpha_0} e^{\alpha_0 \omega^{p-1} s} & (\omega^{p-1} s)^{\alpha_1} e^{\alpha_1 \omega^{p-1} s} & \cdots & (\omega^{p-1} s)^{\alpha_{p-1}} e^{\alpha_{p-1} \omega^{p-1} s} \end{vmatrix} \tag{17}$$

(c'est le déterminant des équations (16)). Le déterminant (17) est de la forme

$$\delta(s) = \begin{vmatrix} f_0(\lambda_0 s) & f_1(\lambda_0 s) & \cdots & f_{p-1}(\lambda_0 s) \\ f_0(\lambda_1 s) & f_1(\lambda_1 s) & \cdots & f_{p-1}(\lambda_1 s) \\ \cdot & \cdot & \cdots & \cdot \\ f_0(\lambda_{p-1} s) & f_1(\lambda_{p-1} s) & \cdots & f_{p-1}(\lambda_{p-1} s) \end{vmatrix}; \tag{18}$$

$\lambda_0, \lambda_1, \cdots, \lambda_{p-1}$ sont des constantes, $f_0(s), f_1(s), \cdots, f_{p-1}(s)$ des fonctions que nous supposons développables en série de Maclaurin. Le déterminant (18) est "composé" de deux rectangles à p lignes et à une infinité de colonnes:

$$\left\| \begin{matrix} f_0(0), & f_0'(0), & f_0''(0), & \cdots, f_0^{(n)}(0), & \cdots \\ \cdot & \cdot & \cdot & \cdots \quad \cdot & \cdots \\ f_{p-1}(0), & f_{p-1}'(0), & f_{p-1}''(0), & \cdots, f_{p-1}^{(n)}(0), & \cdots \end{matrix} \right\|,$$

$$\left\| \begin{matrix} 1, & \frac{\lambda_0 s}{1!}, & \frac{\lambda_0^2 s^2}{2!}, & \cdots, \frac{\lambda_0^n s^n}{n!}, & \cdots \\ \cdot & \cdot & \cdot & \cdots \quad \cdot & \cdots \\ 1, & \frac{\lambda_{p-1} s}{1!}, & \frac{\lambda_{p-1}^2 s^2}{2!}, & \cdots, \frac{\lambda_{p-1}^n s^n}{n!}, & \cdots \end{matrix} \right\|$$

(il est, comme on dit quelquefois, le "produit" de ces deux rectangles). Le déterminant (18) est égal à une somme d'une infinité de termes, chacun étant le produit de deux facteurs; le premier facteur est un déterminant à p lignes, tiré du premier rectangle, le second facteur est le déterminant correspondant tiré du second rectangle. Cherchons parmi tous ces termes celui qui est de degré minimum en s; nous n'écrirons que ce terme (légèrement transformé) pour suggérer tout le développement:

$$\delta(s) = \begin{vmatrix} f_0(0) & f_0'(0) & \cdots & f_0^{(p-1)}(0) \\ \cdot & \cdot & \cdots & \cdot \\ f_{p-1}(0) & f_{p-1}'(0) & \cdots & f_{p-1}^{(p-1)}(0) \end{vmatrix} \cdot \begin{vmatrix} 1 & \lambda_0 & \cdots & \lambda_0^{p-1} \\ \cdot & \cdot & \cdots & \cdot \\ 1 & \lambda_{p-1} & \cdots & \lambda_{p-1}^{p-1} \end{vmatrix} \frac{s^{1+2+\cdots+p-1}}{1!2!\cdots(p-1)!} + \cdots. \tag{19}$$

$D(s)$ est un cas particulier de $\delta(s)$. Pour obtenir ce cas particulier, il faut d'abord poser

$$\lambda_0 = 1, \lambda_1 = \omega, \cdots, \lambda_{p-1} = \omega^{p-1}.$$

Parmi ces p nombres il n'y a pas deux égaux, donc le second déterminant dans le second membre de (19), le Vandermondien de ces p nombres, est $\neq 0$. Puis il faut poser

$$f_0(s) = s^{\alpha_0}e^{a_0 s},\ f_1(s) = s^{\alpha_1}e^{a_1 s},\ \cdots,\ f_{p-1}(s) = s^{\alpha_{p-1}}e^{a_{p-1}s}.$$

Le développement

$$s^{\alpha}e^{as} = \alpha!\frac{s^{\alpha}}{\alpha!} + \frac{(\alpha+1)!}{1!}a\frac{s^{\alpha+1}}{(\alpha+1)!} + \frac{(\alpha+2)!}{2!}a^2\frac{s^{\alpha+2}}{(\alpha+2)!} + \cdots$$

montre que la ligne générale du premier déterminant dans le second membre de (19) a exactement la forme (9); il faut poser

$$\alpha = \alpha_k, \qquad a = a_k$$

et, après cela, $k=0, 1, 2, \cdots, p-1$; le déterminant est donc identique à Δ et, par conséquent, $\neq 0$, voir §2. Notons le résultat:

La fonction entière $D(s)$ pour laquelle le point $s=0$ est un zéro dont l'ordre est exactement $\frac{1}{2}p(p-1)$, n'est pas identiquement nulle.

II. Nous définissons $L_0, L_1, \cdots, L_{p-1}$ par les équations (16). Nous venons de voir que $D(s)$, le déterminant de ces équations, n'est pas identiquement nul; donc

$$L_k = L_k(z, s), \qquad k = 0, 1, 2, \cdots, p-1,$$

est bien défini, sauf aux zéros de $D(s)$. Or L_k est le quotient de deux déterminants; cette expression montre que $L_k(z, s)$ est une *fonction entière de z* (dépendant du paramètre s) et une *fonction méromorphe de s* (dépendant du paramètre z); les pôles de $L_k(z, s)$ (comme fonction de s) se trouvent parmi les zéros de $D(s)$.

Pour la commodité de l'écriture, nous posons

$$\frac{\partial^n L(z, s)}{\partial z^n} = L^{(n)}(z, s).$$

Dérivons l'équation générale du système (16) c.à.d.

$$e^{z\omega^k s} = \sum_{j=0}^{p-1}(\omega^k s)^{\alpha_j}e^{a_j\omega^k s}L_j(z, s)$$

α_l+pm fois par rapport à z et, après cela, posons $z=a_l$. Nous obtenons

$$(\omega^k s)^{\alpha_l}e^{a_l\omega^k s}s^{pm} = \sum_{j=0}^{p-1}(\omega^k s)^{\alpha_j}e^{a_j\omega^k s}L_j^{(\alpha_l+pm)}(a_l, s) \tag{20}$$

pour $k=0, 1, \cdots, p-1$. C'est un système de p équations aux p inconnues

$$L_0^{(\alpha_l+pm)}(a_l, s),\ L_1^{(\alpha_l+pm)}(a_l, s), \cdots, L_{p-1}^{(\alpha_l+pm)}(a_l, s).$$

dont le déterminant est $D(s)$; mais $D(s)$ n'est pas identiquement $=0$; donc les inconnues, comme fonctions méromorphes de s, sont univoquement déterminées. D'autre part on voit immédiatement que les valeurs

$$L_l^{(a_l+pm)}(a_l, s) = s^{pm}, \qquad L_j^{(a_l+pm)}(a_l, s) = 0, \qquad j \neq l, \tag{21}$$

satisfont au système (20). Nous avons ainsi obtenu les relations remarquables (21), valables pour $j, l=0, 1, \cdots, p-1$ et $m=0, 1, 2, \cdots$ (6).

III. Nous pouvons décomposer notre problème, qui est de satisfaire aux équations (1'), en p problèmes un peu plus simples. En effet, fixons k, $0 \leqq k \leqq p-1$, et cherchons une fonction entière $F_k(z)$ qui satisfait aux équations suivantes:

$$F_k^{(a_k+pm)}(a_k) = A_{k+pm}, \qquad F_k^{(a_j+pm)}(a_j) = 0, \qquad j \neq k, \tag{22}$$

$m=0, 1, 2, \cdots$. (Cela revient à laisser les seconds membres tels qu'ils sont dans une certaine colonne du tableau (1') et à les remplacer par 0 dans les $p-1$ autres colonnes.) Ayant trouvé $F_0(z), F_1(z), \cdots, F_{p-1}(z)$,

$$F(z) = F_0(z) + F_1(z) + \cdots + F_{p-1}(z)$$

sera évidemment une solution du système (1').

5. **Fin de la démonstration.** Nous ne perdons rien d'essentiel en nous limitant à la démonstration de l'existence de $F_0(z)$. Simplifions l'écriture et écrivons

$$F(z), \qquad L(z), \qquad a, \qquad 0$$

à la place de

$$F_0(z), \qquad L_0(z), \qquad a_0, \qquad \alpha_0$$

(on sait, d'après (10), que $\alpha_0=0$). Avec ces notations, nous avons

$$F^{(pm)}(a) = A_{pm}, \qquad F^{(a_l+pm)}(a_l) = 0, \tag{23}$$

$$L^{(pm)}(a, s) = s^{pm}, \qquad L^{(a_l+pm)}(a_l, s) = 0 \tag{24}$$

pour $m=0, 1, 2, \cdots$ et $l=1, 2, \cdots, p-1$. Observons que c'est seulement une partie des relations (21), p relations sur p^2, que nous venons d'écrire sous la forme (24). Observons encore qu'en vertu de (6) et de (10)

$$\alpha_{pn} = \alpha_0 + pn = pn. \tag{25}$$

Soit C_n une circonférence de centre $s=0$ qui ne passe par aucun des zéros

(6) On établit de la même manière, en se servant de l'unicité de la solution, la propriété (15). Encore d'autres propriétés de $D(s)$ et de $L(s, s)$ jouent un rôle dans le problème de convergence mentionné (5). P.e. il est facile de démontrer que $D(s)$ a une infinité de racines si $\alpha_0+\alpha_1+\cdots+\alpha_{p-1}<0+1+\cdots+p-1$.

de $D(s)$, donc par aucun pôle de $L(z, s)$. Nous allons trouver une solution $F(z)$ des équations (23) qui est de la forme(7)

$$F(z) = \sum_{n=0}^{\infty} \frac{A_{pn}}{2\pi i} \oint_{C_n} \frac{L(z, s)ds}{s^{pn+1}}. \tag{26}$$

I. Nous admettons d'abord que la série (26) converge uniformement en chaque domaine borné du plan des z.

Alors $F(z)$, donnée par (26), est une fonction entière de z puisque chaque terme de la série est une fonction entière de z.

Puis nous trouvons, en nous servant de (24)

$$F^{(pm)}(a) = \sum_{n=0}^{\infty} \frac{A_{pn}}{2\pi i} \oint_{C_n} \frac{s^{pm}ds}{s^{pn+1}} = A_{pm},$$

$$F^{(al+pm)}(a_l) = \sum_{n=0}^{\infty} \frac{A_{pn}}{2\pi i} \oint_{C_n} \frac{0ds}{s^{pn+1}} = 0, \qquad l > 0.$$

c.à.d. les équations (23) sont satisfaites.

II. Que devons-nous démontrer encore? La convergence uniforme de la série (26) pour

$$|z| \leqq R \tag{27}$$

où R est un nombre positif donné (arbitrairement grand). De quoi pouvons-nous disposer encore? Des cercles C_n.

Je désignerai dans ce qui suit par $k_1, k_2, \cdots, K, k$ des nombres positifs qui dépendent des données du problème (de $a_0, a_1, \cdots, a_{p-1}, \alpha_0, \alpha_1, \cdots, \alpha_{p-1}$) ainsi que de R, mais qui sont indépendants de s et de n. La définition (17) de $D(s)$ montre qu'il existe k_1 et k_2 tels que pour tout s

$$|D(s)| < k_1 e^{k_2|s|}.$$

Vu cette inégalité, nous pouvons, d'après un théorème de Valiron(8), choisir r_n tel que

$$pn < r_n < 2pn \tag{28}$$

et que, en chaque point s du cercle $|s| = r_n$,

(7) En cherchant une solution de la forme (26), je me sers d'une méthode que *Hurwitz* a employée pour un but différent; voir Acta Mathematica, Tome 20 (1897), pp. 285–312 ou *Mathematische Werke*, Basel, 1932, Tome 1, pp. 436–460. Dans une Thèse, faite sous ma direction, *H. Muggli* a appliqué la méthode de Hurwitz à des problèmes assez généraux; voir Commentarii Mathematici Helvetici, Tome 11 (1938), pp. 151–179.

(8) *G. Valiron, Integral Functions*, Toulouse, 1923; voir le Théorème 27, p. 89. Dans la Thèse citée (7), la généralité des applications résulte également de l'usage de ce théorème de Valiron.

$$|D(s)| > \frac{1}{k_3 e^{k_4|s|}}, \tag{29}$$

cette inégalité étant valable pour $n=1, 2, 3, \cdots$. *Soit r_n le rayon de C_n*; donc (29) est valable sur le cercle C_n.

$L(z, s)$ est un quotient des deux déterminants, d'après (16). La forme du numérateur montre que, z étant un point quelconque du cercle (27), on a pour tout s

$$|D(s)L(s, z)| < k_5 e^{k_6|s|};$$

en combinant cette inégalité avec (29) on obtient

$$|L(z, s)| < k_5 e^{k_6|s|} \cdot k_3 e^{k_4|s|} = K e^{k|s|} \tag{30}$$

sur tous les cercles C_n, $n=1, 2, 3, \cdots$.

A l'aide de (28) et de (30) nous évaluons le module du terme général de la série (26) dans le cercle (27):

$$\left|\frac{A_{pn}}{2\pi i}\oint_{C_n}\frac{L(z, s)ds}{s^{pn+1}}\right| < \frac{|A_{pn}|}{2\pi}\,\frac{Ke^{kr_n}2\pi r_n}{r_n^{pn+1}} < |A_{pn}|\,K\frac{e^{k2pn}}{(pn)^{pn}} \leqq \frac{|A_{pn}|}{(pn)!}\,Ke^{k2pn}. \tag{31}$$

La n-ième racine de cette dernière quantité tend, en vertu de l'hypothèse (4), vers 0 avec $1/n$; il faut tenir compte de (25). Donc, la série (26) possède dans le cercle (27) une majorante numérique convergente. C.q.f.d.

6. **Problèmes de la même forme, qui ne possèdent pas de solution.** Nous considérons le problème de trouver une fonction entière qui satisfait à (1). Nous admettons les hypothèses (2), (3), (4) et en outre la suivante (qui figure aussi dans le théorème que nous venons de démontrer): Chaque terme de la suite infinie $a_0, a_1, a_2, \cdots$ est égal à un certain terme d'une suite finie donnée. Ces hypothèses sont insuffisantes pour assurer l'existence de la solution comme les exemples suivants le montrent.

1er exemple. La suite finie contient trois valeurs différentes c, 0, $-c$. Les trois suites infinies sont définies ainsi:

$$\begin{aligned} a_n &= c, -c; \quad 0, \quad 0, \quad 0, \quad 0, \cdots; \\ \alpha_n &= 0, \quad 0; \quad 1, \quad 3, \quad 5, \quad 7, \cdots; \\ A_n &= 0, \quad 1; \quad 0, \quad 0, \quad 0, \quad 0, \cdots. \end{aligned}$$

La loi des conditions successives est uniforme à partir de la 3me condition (après le signe ;) et, d'après cette partie des conditions, $F(z)$ devrait être une fonction paire—mais cela contredit les deux premières conditions; $F(z)$ n'existe pas.

Cet exemple est plutôt trivial, mais montre clairement ce qu'il doit montrer: l'insuffisance des hypothèses énoncées pour assurer l'existence de la solution. L'exemple suivant est moins trivial.

2me exemple. La suite finie ne contient que deux termes différents: a et b. Les trois suites infinies sont définies ainsi:

$$a_n = a,\ b;\ a,\ b,\ a,\ a,\ b;\ a,\ b,\ a,\ a,\ b;\cdots;$$
$$\alpha_n = 0,\ 0;\ 1,\ 1,\ 2,\ 3,\ 3;\ 5,\ 5,\ 6,\ 7,\ 7;\cdots;$$
$$A_n = 0,\ 1;\ 0,\ 0,\ 0,\ 0,\ 0;\ 0,\ 0,\ 0,\ 0,\ 0;\cdots.$$

La loi des conditions successives est uniforme à partir de la 3me condition (après le premier signe ;). On a pour $n \geqq 3$

$$a_{n+5} = a_n, \qquad \alpha_{n+5} = \alpha_n + 4, \qquad A_n = 0.$$

D'après cette partie des conditions, $F(z)$ devrait être paire relativement aux deux points a et b et avoir une symétrie encore plus haute relativement au point a; plus précisément, $F(z)$ devrait satisfaire aux identiés par rapport à z:

$$F(z) = F(a + i(z - a)), \qquad F(z) = F(b - (z - b)).$$

De ces identités on tire, par calcul formel ou en tenant compte de la signification géométrique:

$$F(z + 2(b - a)) = F(z), \qquad F(z + 2i(b - a)) = F(z).$$

$F(z)$ serait ainsi une fonction entière doublement périodique, donc une constante—mais cela contredit les deux premières conditions; $F(z)$ n'existe pas.

ÉCOLE POLYTECHNIQUE FÉDÉRALE,
ZURICH, SUISSE,
BROWN UNIVERSITY,
PROVIDENCE, R. I.

THE PARASTROPHIC CRITERION FOR THE FACTORIZATION OF PRIMES

BY

FRED KIOKEMEISTER

If the rational prime p is not an extraordinary divisor of the discriminant of the irreducible polynomial $f(x)$ with coefficients in the domain of rational integers, the factorization of $f(x)$ modulo p gives the factorization of the principal ideal (p) in the maximal domain of integrity of the algebraic number field defined by $f(x)$. This criterion was given by Zolotareff[1] and discussed by Dedekind[2]. Hilbert, by relating the structure of (p) to the factorization modulo p of the minimal polynomial of the general element of the maximal domain of integrity[3], removed the necessity of separately considering the divisors of the discriminant. The parastrophic criterion here presented serves the same purpose with what is perhaps a method involving a simpler computation.

Parts 2 and 4 of the paper are related to the work of Nakayama and Nesbitt[4] on the representations of Frobenius algebras and to the study of such algebras by Nakayama[5].

1. **The parastrophic matrix.** Let A be a linear associative algebra with basis $e_1, e_2, \cdots, e_n$ over a field P. We designate the vectors of basis elements as follows:

$$u = (e_1, e_2, \cdots, e_n), \qquad u' = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix},$$

so that the vector u' is the transpose of u. We shall be consistent in employing this notation for a vector and its transpose throughout the paper. From the theory[6] of the matric representations of a linear algebra, we have for an element a contained in A.

$$au = uR(a), \qquad u'a = S(a)u', \tag{1.1}$$

Presented to the Society April 12, 1941; received by the editors July 30, 1940.

(1) G. Zolotareff, *Théories des Nombres Entiers Complexes, avec une Application au Calcul Intégral*, St. Petersburg, 1874.

(2) Dedekind, Göttingen Abhandlungen, vol. 23 (1878).

(3) Hilbert, *Die Theorie der algebraischen Zahlkörper*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 4 (1894–1895), p. 195.

(4) Nakayama and Nesbitt, Annals of Mathematics, (2), vol. 39 (1938), pp. 659–668.

(5) Nakayama, Annals of Mathematics, (2), vol. 40 (1939), pp. 611–633.

(6) M. Deuring, *Algebren*, Berlin, 1935, p. 2.

where $R(a)$ and $S(a)$ are, respectively, the first and second representations of a. The elements of the vectors au and $u'a$ are linear forms in the basis elements $e_1, e_2, \cdots, e_n$ with coefficients in P; the elements of the matrices $R(a)$ and $S(a)$ are uniquely determined in P.

The multiplication table of the algebra A can be given as the outer product of the vectors u' and u:

$$u'u = \begin{bmatrix} e_1^2 & e_1e_2 \cdots e_1e_n \\ e_2e_1 & e_2^2 \cdots e_2e_n \\ \cdot & \cdot \cdots \cdot \\ \cdot & \cdot \cdots \cdot \\ e_ne_1 & \cdot \cdots e_n^2 \end{bmatrix} = \begin{bmatrix} e_1u \\ e_2u \\ \cdot \\ \cdot \\ e_nu \end{bmatrix} = (u'e_1, u'e_2, \cdots, u'e_n).$$

Employing equations (1.1), we have immediately

$$u'u = \begin{bmatrix} uR_1 \\ uR_2 \\ \cdot \\ \cdot \\ uR_n \end{bmatrix} = (S_1u', S_2u', \cdots, S_nu') \tag{1.2}$$

where $R_i = R(e_i)$, $S_i = S(e_i)$. The matrix $u'u$ has as elements linear forms in the basis elements. We shall designate $u'u$ by $Q(u)$. The array $Q(u)$ is uniquely defined by the basis u employed in the representation of the algebra A. The algebra is determined up to an isomorphic ring and isomorphic basis by $Q(u)$.

Frobenius defined the parastrophic matrix(7) of an algebra A to be that derived from $Q(u)$ by replacing each basis element e_i by the variable x_i. This replacement may be achieved by the substitution of $\xi = (x_1, x_2, \cdots, x_n)$ for u and ξ' for u' in the forms (1.2). In this notation the parastrophic matrix is

$$Q(\xi) = \begin{bmatrix} \xi R_1 \\ \xi R_2 \\ \cdot \\ \cdot \\ \xi R_n \end{bmatrix} = (S_1\xi', S_2\xi', \cdots, S_n\xi'). \tag{1.3}$$

The variables $x_1, x_2, \cdots, x_n$ range over P.

THEOREM 1.1. *If* $\xi = (x_1, x_2, \cdots, x_n)$, $\eta = (y_1, y_2, \cdots, y_n)$ *and if* α_1, α_2 *are elements of* P,

$$Q(\alpha_1\xi + \alpha_2\eta) = \alpha_1Q(\xi) + \alpha_2Q(\eta).$$

The theorem follows from equation (1.3); for

(7) Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1903, p. 507.

$$Q(\alpha_1\xi+\alpha_2\eta)=\begin{bmatrix}(\alpha_1\xi+\alpha_2\eta)R_1\\(\alpha_1\xi+\alpha_2\eta)R_2\\\vdots\\(\alpha_1\xi+\alpha_2\eta)R_n\end{bmatrix}=\alpha_1\begin{bmatrix}\xi R_1\\\xi R_2\\\vdots\\\xi R_n\end{bmatrix}+\alpha_2\begin{bmatrix}\eta R_1\\\eta R_2\\\vdots\\\eta R_n\end{bmatrix}=\alpha_1Q(\xi)+\alpha_2Q(\eta).$$

Let $a=\alpha_1e_1+\alpha_2e_2+\cdots+\alpha_ne_n$, $\alpha_i\in P$, be an element of A. Then $\bar{a}=(\alpha_1, \alpha_2, \cdots, \alpha_n)$ is the unique vector representation of a with respect to the basis u. Employing the inner product of vectors, we have

$$a=\bar{a}u'=u\bar{a}'. \tag{1.4}$$

The vector representation of the elements of the algebra A has the following properties:

THEOREM 1.2. *For any elements a and b in A,*

(a) $\overline{\alpha a}=\alpha\bar{a}$, $\alpha\in P$.

(b) $\overline{a+b}=\bar{a}+\bar{b}$.

(c) $(\overline{ab})'=R(a)\bar{b}'$, $\overline{ab}=\bar{a}S(b)$.

Parts (a) and (b) of the theorem follow from the definition of an algebra.

For the proof of part (c), let a and b be two elements of A. Consider that by equation (1.4)

$$ab=a(u\bar{b}')=(au)\bar{b}'.$$

Then by (1.1)

$$ab=(uR(a))\bar{b}'=u(R(a)\bar{b}').$$

It follows from the uniqueness of the vector representation that

$$(\overline{ab})'=R(a)\bar{b}'.$$

Similarly

$$\overline{ab}=\bar{a}S(b).$$

THEOREM 1.3. *For any element a of the algebra A, the relations hold:*

$$\bar{a}Q(\xi)=\xi R(a), \qquad Q(\xi)\bar{a}'=S(a)\xi'$$

where $\xi=(x_1, x_2, \cdots, x_n)$, and $Q(\xi)$ is the parastrophic matrix of the fixed basis u.

Since $\bar{a}u'=a$,

$$\bar{a}u'u=\bar{a}Q(u)=au=uR(a),$$

so that

$$\bar{a}Q(u)=uR(a).$$

Here we have the equality of two vectors with elements which are linear forms

in the basis elements $e_1, e_2, \cdots, e_n$. The basis elements are linearly independent, and the linear forms must be identical. Since $\bar{a}$ and $R(a)$ have elements in the field P, we may replace $u=(e_1, e_2, \cdots, e_n)$ by $\xi=(x_1, x_2, \cdots, x_n)$:

$$\bar{a}Q(\xi) = \xi R(a).$$

Similarly $u\bar{a}'=a$,

$$u'u\bar{a}' = Q(u)\bar{a}' = u'a = S(a)u',$$

and

$$Q(u)\bar{a}' = S(a)u'.$$

This gives

$$Q(\xi)\bar{a}' = S(a)\xi'$$

by the same argument as above.

An algebra A is said to be a *Frobenius algebra* when there exists a vector $\eta=(\delta_1, \delta_2, \cdots, \delta_n)$ with elements δ_i in P such that $Q(\eta)$ is nonsingular.

2. **Orthogonal modules.** In this section we shall assume that the algebra A has an identity e. Any linear algebra not a null algebra is homomorphic to an algebra with identity.

LEMMA 2.1. *If A has an identity e, $Q(\eta)=0$ implies that $\eta=(0, 0, \cdots, 0)$.*

Let Theorem 1.3 be applied to the element e and the vector η where $Q(\eta)=0$:

$$\bar{e}Q(\eta) = \eta R(e) = \eta.$$

Then if $Q(\eta)=0$, $\eta=0$.

Let M designate the set of all vectors $\eta=(\delta_1, \delta_2, \cdots, \delta_n)$ where $\delta_1, \delta_2, \cdots, \delta_n$ are elements of P, i.e., M is the set of all constant vectors when the variables x_i of $\xi=(x_1, x_2, \cdots, x_n)$ are allowed to run over P.

LEMMA 2.2. *If D is a set of elements d in A, the set $U(D)$ of all $\eta\in M$ such that for all $d\in D$*

$$\bar{d}Q(\eta) = 0$$

forms a P-module closed on the right under multiplication by all $S^{\tau}(x)$, $x\in A$. The set $V(D)$ of all $\eta\in M$ such that for all $d\in D$

$$Q(\eta)\bar{d}' = 0$$

forms a P-module closed on the right under multiplication by $R(x)$, $x\in A$.

Let η_1, η_2 be such that

$$\bar{d}Q(\eta_i) = 0, \qquad d \in D.$$

Then for all $d \in D$, by Theorem 1.1,

$$\bar{d}Q(\alpha_1\eta_1 + \alpha_2\eta_2) = \alpha_1\bar{d}Q(\eta_1) + \alpha_2\bar{d}Q(\eta_2) = 0$$

and $U(D)$ includes all linear forms in η_1, η_2 with coefficients in P, i.e., $U(D)$ is a P-module. Furthermore, since, for the general vector ξ, $\bar{d}Q(\xi) = \xi R(d)$ (Theorem 1.3), $\bar{d}Q(\eta) = 0$ implies that

$$\bar{d}Q(\eta S^\tau(x)) = \eta S^\tau(x)R(d) = \eta R(d)S^\tau(x) = 0$$

for all $d \in D$, and $x \in A$ by the commutivity of $R(d)$ with $S^\tau(x)$ as established by Frobenius([8]). It follows that $\eta S^\tau(x)$ is an element of $U(D)$ whenever η is, and $U(D)$ is closed on the right under multiplication by all $S^\tau(x)$.

The second part of the theorem is proved by applying the same argument to the relation $Q(\xi)\bar{d}' = S(d)\xi'$.

For any theorem which is proved for right ideals in the following pages there exists an analogous theorem for left ideals. Since the argument in each case is the same, we shall state the theorems in most instances only for right ideals.

LEMMA 2.3. *If N is a subset of the module M, the set $\mathfrak{a}$ of elements $a \in A$ such that*

$$\bar{a}Q(\eta) = 0$$

for all $\eta \in N$ forms a right ideal in A. Furthermore, $U(\mathfrak{a}) \subseteq N$.

Let a_1 and a_2 be elements of $\mathfrak{a}$. Then for all $\eta \in N$

$$\overline{(\alpha_1 a_1 + \alpha_2 a_2)}Q(\eta) = (\alpha_1\bar{a}_1 + \alpha_2\bar{a}_2)Q(\eta) = \alpha_1\bar{a}_1Q(\eta) + \alpha_2\bar{a}_2Q(\eta) = 0$$

where α_1, α_2 are elements of P, and $\mathfrak{a}$ is a P-module. The set $\mathfrak{a}$ is a right ideal; for let a be an element of $\mathfrak{a}$, c an element of A. Then by Lemma 1.2

$$\overline{ac} = \bar{a}S(c).$$

Frobenius established that

$$S(c)Q(\eta) = Q(\eta)R(c).$$

Employing these results, we have

$$\overline{ac}Q(\eta) = \bar{a}S(c)Q(\eta) = \bar{a}Q(\eta)R(c) = 0$$

for all c in A. Hence ac lies in $\mathfrak{a}$, and $\mathfrak{a}$ is a right ideal. By the definition of $U(\mathfrak{a})$, this set must include N.

The order of an ideal $\mathfrak{a}$ is the order with respect to the basis field P; $\mathfrak{a}$ is of order r if r linearly independent elements constitute a P-basis for $\mathfrak{a}$.

([8]) Frobenius, ibid., p. 507.

LEMMA 2.4. *If $\mathfrak{a}$ is a right ideal of order r, $U(\mathfrak{a})$ is a P-module of order $n-r$.*

If $\mathfrak{a}=A$, then $r=n$. As in the proof of Lemma 2.1, $\bar{e}Q(\eta)=0$ implies that η is the zero vector. Then $U(\mathfrak{a})=0$, and the theorem holds.

If $\mathfrak{a}=0$, then $r=0$. Let $\bar{0}$ designate the zero vector which represents the element 0 of the algebra. Every vector η satisfies the equation $\bar{0}Q(\eta)=0$, so that $U(\mathfrak{a})$ is of order n. In this special case also the theorem holds.

If $\mathfrak{a}\neq A$ and $\mathfrak{a}\neq 0$, then $\mathfrak{a}$ has a basis $v=(a_1, a_2, \cdots, a_r)$. Let $s_v=(e_1, e_2, \cdots, e_{n-r})$ be a set of elements supplementary to v so that $u=(s_v, v)$ is a basis of A. Since $\mathfrak{a}\neq A$, e is not included in $\mathfrak{a}$, so we may choose e_1 equal to the identity element of A.

If we employ $u=(s_v, v)$ to form $Q(u)$, we have

$$Q(u) = u'u = (s_v, v)'(s_v, v) = \begin{bmatrix} s_v's_v & s_v'v \\ v's_v & v'v \end{bmatrix}.$$

Since $\mathfrak{a}$ is a right ideal, every element of $v's_v$ and $v'v$ is expressible in terms of the basis v. Writing $Q(u)$ as a matrix of single elements, we have

$$Q(u) = \begin{bmatrix} e_1, & e_2, \cdots, e_{n-r}, & a_1, \cdots, a_r \\ e_2, & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ e_{n-r}, & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ a_1, & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ a_r, & \cdot & \cdot \quad \cdot \quad \cdot \end{bmatrix}$$

where every element of $Q(u)$ in any row below the $(n-r)$th row must be expressed in the basis v. We pass by substitution of $\xi=(x_1, x_2, \cdots, x_n)$ for the basis vector u to the parastrophic matrix $Q(\xi)$ and then set $x_i=\alpha_i\in P$, $\eta=(\alpha_1, \alpha_2, \cdots, \alpha_n)$:

$$Q(\eta) = \begin{bmatrix} \alpha_1, & \alpha_2, \cdots, \alpha_{n-r}, & \alpha_{n-r+1}, \cdots, \alpha_n \\ \alpha_2, & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \alpha_{n-r}, & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \alpha_{n-r+1}, & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \alpha_n, & \cdot & \cdot \quad \cdot \quad \cdot \end{bmatrix}$$

where the elements of any row below the $(n-r)$th are linear forms in $\alpha_{n-r+1}, \cdots, \alpha_n$. Now if $a \in \mathfrak{a}$,

$$\bar{a} = (0, \cdots, 0, \beta_{n-r+1}, \cdots, \beta_n) \qquad \text{where } \beta_i \in P.$$

Obviously if we choose $\alpha_j = 0$ for $j = n-r+1, \cdots, n$, then

$$\bar{a}Q(\eta) = 0$$

for any choice of $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$. For in this case every element below the $(n-r)$th row is zero. Conversely if $\bar{a}Q(\eta) = 0$ for all $a \in \mathfrak{a}$, then in particular

$$\beta_{n-r+1}\alpha_{n-r+1} + \cdots + \beta_n\alpha_n = 0$$

for all $\beta_j \in P$. This linear form is identically zero only if the α_j involved are all zero.

It follows that $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_{n-r}, 0, \cdots, 0)$ is the most general vector such that

$$\bar{a}Q(\eta) = 0$$

for all $a \in \mathfrak{a}$. But $U(\mathfrak{a})$ is this set of vectors, and is therefore of order $n-r$.

If there exists a vector η such that $\eta R(a) = 0$, η is said to be *orthogonal on the left* to $R(a)$. Let a be an element of the right ideal $\mathfrak{a}$. From the relation

$$\bar{a}Q(\eta) = \eta R(a),$$

we see that $\eta R(a) = 0$ for all $a \in \mathfrak{a}$ implies that $\bar{a}Q(\eta) = 0$ and conversely. That is, η must lie in $U(\mathfrak{a})$. Similarly if $\mathfrak{b}$ is a left ideal, $V(\mathfrak{b})$ is the set of all vectors such that

$$Q(\eta)\bar{b}' = S(b)\eta' = 0,$$

and η is *orthogonal on the right* to all $S(b)$ where b lies in $\mathfrak{b}$.

THEOREM 2.1. *If $\mathfrak{a}$ is a right ideal, then $U(\mathfrak{a})$ is the set of vectors orthogonal on the left to all $R(a)$, $a \in \mathfrak{a}$. If $\mathfrak{b}$ is a left ideal, $V(\mathfrak{b})$ is the set of all vectors orthogonal on the right to all $S(b)$, $b \in \mathfrak{b}$.*

The vector sets $U(\mathfrak{a})$ and $V(\mathfrak{b})$ are called the *orthogonal sets* of the right ideal $\mathfrak{a}$ and the left ideal $\mathfrak{b}$, respectively.

LEMMA 2.5. (a) $U(a) = 0$, $U(0) = M$. *If* $U_i = U(\mathfrak{a}_i)$, *where* $\mathfrak{a}_i$ *is a right ideal,* (b) $\mathfrak{a}_1 \supset \mathfrak{a}_2$ *if and only if* $U_1 \subset U_2$, (c) $U(\mathfrak{a}_1, \mathfrak{a}_2) = U_1 \wedge U_2$, $U(\mathfrak{a}_1 \wedge \mathfrak{a}_2) = (U_1, U_2)$.

Part (a) follows immediately from Lemma 2.4.

Part (b): Let $a_i \in \mathfrak{a}_i$, $\eta_i \in U_i$. Then

$$\bar{a}_1 Q(\eta_1) = 0$$

for all a_1 in $\mathfrak{a}_1$. In particular

$$a_2 \in \mathfrak{a}_2 \subset \mathfrak{a}_1,$$

and hence

$$\bar{a}_2 Q(\eta_1) = 0$$

so that all η_1 are elements of U_2, or $U_1 \subset U_2$.

Part (c): If $\eta \in U_1 \wedge U_2$, then $\bar{a}Q(\eta) = 0$ for all $a \in (\mathfrak{a}_1, \mathfrak{a}_2)$ so that

$$U_1 \wedge U_2 \subseteq U(\mathfrak{a}_1, \mathfrak{a}_2).$$

By (b) above, and since $\mathfrak{a}_i \subset (\mathfrak{a}_1, \mathfrak{a}_2)$,

$$U_i \supset U(\mathfrak{a}_1, \mathfrak{a}_2), \qquad U_1 \wedge U_2 \supseteq U(\mathfrak{a}_1, \mathfrak{a}_2).$$

Finally these two relations together imply that

$$U_1 \wedge U_2 = U(\mathfrak{a}_1, \mathfrak{a}_2).$$

Similarly, let $\eta \in (U_1, U_2)$. Then $\bar{a}Q(\eta) = 0$ for all $a \in (\mathfrak{a}_1 \wedge \mathfrak{a}_2)$ so that

$$(U_1, U_2) \subseteq (\mathfrak{a}_1 \wedge \mathfrak{a}_2).$$

However, if $\mathfrak{a}_i$ is of order r_i, by Lemma 2.4

$$\begin{aligned} \text{order } (U_1, U_2) &= \text{order } U_1 + \text{order } U_2 - \text{order } (U_1 \wedge U_2) \\ &= n - r_1 + n - r_2 - \text{order } (U_1 \wedge U_2), \end{aligned}$$

From the equation $(U_1 \wedge U_2) = U(\mathfrak{a}_1, \mathfrak{a}_2)$ we have

$$\text{order } (U_1 \wedge U_2) = n - \text{order } (\mathfrak{a}_1, \mathfrak{a}_2),$$

and we know that

$$\text{order } (\mathfrak{a}_1, \mathfrak{a}_2) = r_1 + r_2 - \text{order } (\mathfrak{a}_1 \wedge \mathfrak{a}_2),$$

so that

$$\text{order } (U_1, U_2) = n - \text{order } (\mathfrak{a}_1 \wedge \mathfrak{a}_2) = \text{order } U(\mathfrak{a}_1 \wedge \mathfrak{a}_2).$$

We have proved that $U(\mathfrak{a}_1 \wedge \mathfrak{a}_2)$ contains (U_1, U_2) and that these two modules have the same order. It must be true that

$$U(\mathfrak{a}_1 \wedge \mathfrak{a}_2) = (U_1, U_2).$$

LEMMA 2.6. *Let η be a vector with elements in P, and let $Q(\eta)$ be of rank r. There exist a right ideal $\mathfrak{a}$ and a left ideal $\mathfrak{b}$, each of order $n-r$, such that $U(\mathfrak{a})$ is the set of vectors $\eta S^{\tau}(x)$, $V(\mathfrak{b})$ is the set of vectors $\eta R(x)$ where x ranges over all elements of the algebra A.*

Let $\mathfrak{a}$ be the set of elements of A whose representation vectors are orthogonal to $Q(\eta)$ on the left. Since $Q(\eta)$ is of rank r, there must be $n-r$ linearly independent such vectors. By Lemma 2.3, $\mathfrak{a}$ is a right ideal in A. Similarly

there will exist the left ideal $\mathfrak{b}$ of order $n-r$ whose elements b are those for which $Q(\eta)b'=0$.

By Lemma 2.2, $\eta S^{\tau}(x)\in U(\mathfrak{a})$ for every $x\in A$. If $r=0$, $\eta=(0, 0, \cdots, 0)$ by Lemma 2.1; for $Q(\eta)=0$ implies that $\eta=(0, 0, \cdots, 0)$. Then $\mathfrak{a}=\mathfrak{b}=A$, and the theorem is trivial.

If $r\neq 0$, $r\neq n$, choose a basis $b_1, b_2, \cdots, b_{n-r}$ for $\mathfrak{b}$ and let $c_1, c_2, \cdots, c_r$ be supplementary to it in A so that $u=(b_1, b_2, \cdots, b_{n-r}, c_1, c_2, \cdots, c_r)$ is a basis for A. If $r=n$, then $\mathfrak{b}=0$, and $c_1, c_2, \cdots, c_n$ is chosen as a basis for A. In either case, if $c=\sum\alpha_i c_i$, $\alpha_i\in P$, is an element of $\mathfrak{b}$, then $c=0$; for in the first case the c_i are linearly independent of the b_i and in the second case $\mathfrak{b}$ consists of the element 0 alone.

The r vectors

$$\eta S^{\tau}(c_1), \eta S^{\tau}(c_2), \cdots, \eta S^{\tau}(c_r)$$

are linearly independent; for if $c=\sum\gamma_i c_i$, then $\sum\gamma_i(\eta S^{\tau}(c_i))=\eta S^{\tau}(c)=0$ implies that $S(c)\eta'=Q(\eta)\bar{c}'=0$ which implies that c is an element of $\mathfrak{b}$. Then $c=0$, and all γ_i are zero. We have established that the set of vectors $\eta S^{\tau}(x)$ is at least of order r. However by Lemma 2.2 this set is contained in $U(\mathfrak{a})$, and by Lemma 2.4 $U(\mathfrak{a})$ is of order r. Therefore $U(\mathfrak{a})=\{\eta S^{\tau}(x)\}$.

A similar proof gives $V(\mathfrak{b})=\{\eta R(x)\}$ of order r.

THEOREM 2.2. *There is a bi-unique correspondence between right ideals of A and sub-modules U of M such that $\eta\in U$ implies that $\eta S^{\tau}(x)\in U$ for all $x\in A$. There is a bi-unique correspondence between left ideals of A and sub-modules V of M such that $\eta\in V$ implies that $\eta R(x)\in V$ for every $x\in A$.*

Lemmas 2.2 and 2.4 give the existence of $U=U(\mathfrak{a})$ for every right ideal $\mathfrak{a}$; $U(\mathfrak{a})$ is of order $n-r$ if $\mathfrak{a}$ is of order r.

Conversely let U be of order r and closed on the right under multiplication by all $S^{\tau}(x)$. Let $\eta_1, \eta_2, \cdots, \eta_r$ be a basis for U. By Lemma 2.6 there exist right ideals $\mathfrak{a}_i$ with the orthogonal modules $U_i=\{\eta S^{\tau}(x)\}$. Now

$$U=(U_1, U_2, \cdots, U_r).$$

From Lemma 2.5, part (c) it follows that $U=U(\mathfrak{a})$ where $\mathfrak{a}=(\mathfrak{a}_1\wedge\mathfrak{a}_2\wedge\cdots\wedge\mathfrak{a}_r)$. For each such set U there exists a right ideal $\mathfrak{a}$ and $U=U(\mathfrak{a})$. This establishes the theorem for right ideals. The proofs for left ideals can be carried through in the same way.

3. **The parastrophic form.** In the preceding section we have established the fact that the ideals of an algebra with identity e can be put into correspondence with certain sets of vectors $\eta=(\delta_1, \delta_2, \cdots, \delta_n)$, $\delta_i\in P$, i.e., with certain sets of values of $x_1, x_2, \cdots, x_n$ considered as variables over the field P. Let A be restricted to be a Frobenius algebra. This implies the existence of the identity e. Then we may define

$$\pi(\xi) = \pi(x_1, x_2, \cdots, x_n) = |Q(\xi)|$$

as the *parastrophic form* with respect to the basis $e_1, e_2, \cdots, e_n$. Since A is a Frobenius algebra, this polynomial is not identically zero. Since the elements of $Q(\xi)$ are homogeneous linear forms in $x_1, x_2, \cdots, x_n$, $\pi(\xi)$ is homogeneous of degree n in the variables.

LEMMA 3.1. *Every vector solution $\xi=\eta$ of the equation $\pi(\xi)=0$ lies in the orthogonal module of a nonzero ideal of A. Conversely if η lies in the orthogonal module of a nonzero ideal in A, $\pi(\eta)=0$.*

If $\xi=\eta$ is a solution of $\pi(\xi)=0$, then $\pi(\eta)=0$, $|Q(\eta)|=0$, and $Q(\eta)$ is of rank less than n. By Lemma 2.6 there exist both left and right ideals whose orthogonal modules contain η.

If $\mathfrak{a}$ is a right ideal of order $r\neq 0$, then by Lemma 2.4 $U(\mathfrak{a})$ is of order $n-r\neq n$, and $a\in\mathfrak{a}$ implies that

$$aQ(\eta) = 0$$

for every η included in $U(\mathfrak{a})$. Then, for every η, $Q(\eta)$ is singular,

$$\pi(\eta) = |Q(\eta)| = 0,$$

and $\xi=\eta$ is a solution of $\pi(\xi)=0$.

THEOREM 3.1. *If A contains a right ideal $\mathfrak{a}$ or a left ideal $\mathfrak{b}$ of order 1, then*

$$\pi(\xi) = g(\xi)h(\xi)$$

where g is a linear form in $x_1, x_2, \cdots, x_n$. The orthogonal module $U(\mathfrak{a})$ or $V(\mathfrak{b})$ consists of all solutions of $g(\xi)=0$.

Let $\mathfrak{a}$ be a right ideal of order 1 in A. Then $U(\mathfrak{a})$ is of order $n-1$, i.e., $U(\mathfrak{a})$ is a linear sub-space of M defined by a linear form $g(\xi)=0$. However since every solution of $g(\xi)=0$ is a solution of $\pi(\xi)=0$, $g(\xi)$ divides $\pi(\xi)$([9]).

THEOREM 3.2. *If A is a division algebra, $\pi(\xi)=0$ can have no solutions other than the zero vector. Furthermore, $\pi(\xi)$ must be of degree n in each of the variables $x_1, x_2, \cdots, x_n$.*

If A is a division algebra, $\pi(\xi)=0$ can have no solution $\eta\neq 0$ since by Lemma 3.1 a zero of $\pi(\xi)$ gives rise to an ideal other than the zero ideal in A. Suppose $\pi(\xi)$ did not contain a term $\alpha_i x_i^n$, $\alpha_i\neq 0$. Then $(0, 0, \cdots, 1, \cdots, 0)$ with 1 in the ith place and 0 elsewhere would be a solution of $\pi(\xi)=0$, by the homogeneity of $\pi(\xi)$.

4. **Commutative Frobenius algebras over a perfect field.** Let A be a commutative algebra of order n over the field P with basis $e_1, e_2, \cdots, e_n$. Every ideal $\mathfrak{a}$ is two sided, and $U(\mathfrak{a})$ is the corresponding orthogonal set.

([9]) Cf. van der Waerden, B. L., *Moderne Algebra*, Berlin, Springer, 1931, vol. 2, p. 11.

LEMMA 4.1. *If A is a commutative Frobenius algebra of order n over P, the parastrophic form $\pi(\xi)$ with respect to the basis $e_1, e_2, \cdots, e_n$ is factorable into linear factors over some finite extension field Δ of P.*

If $n=1$, the theorem is trivial.

Let $n>1$, and consider A/Ω defined with basis $e_1, e_2, \cdots, e_n$ over the algebraic closure Ω of P. Since the same basis is employed, the same multiplication table results and therefore the same parastrophic form.

We define a *minimal ideal* of A to be an ideal which is not the zero ideal and which contains only the zero ideal properly. The representations of A/Ω with respect to its minimal ideals (i.e., the absolutely irreducible representations) are of order 1([10]). Then the minimal ideals themselves are of order 1.

By Theorem 3.1 each minimal ideal induces a linear factor of the parastrophic form $\pi(\xi)$. There can be, therefore, not more than n linear factors associated with the minimal ideals. By Theorem 2.5 these linear factors are distinct if and only if the minimal ideals are distinct. It follows that there can be not more than n minimal ideals. Let $g_1, g_2, \cdots, g_k$ be associated with the minimal ideals $\mathfrak{a}_1, \mathfrak{a}_2, \cdots, \mathfrak{a}_k$ where $k \leqq n$. It is true, by Theorem 3.1, that $\pi(\xi)=g_1^{\sigma_1}g_2^{\sigma_2}\cdots g_k^{\sigma_k}h$. Let $\xi=\eta$ be a solution of the equation $\pi(\xi)=0$. By Lemma 3.1 η must lie in the orthogonal module of some ideal in A. This ideal must contain a minimal ideal $\mathfrak{a}_i$ and therefore η must be a solution of $g_i=0$. Every solution of the equation $\pi(\xi)=0$ must be a solution of the equation $g_1g_2\cdots g_k=0$. We have immediately, by Hilbert's *Nullstellensatz*([11]) that the only factors of $\pi(\xi)$ are products of powers of the linear factors $g_1, g_2, \cdots, g_k$.

If Δ is a field defined through extending the field P by the coefficients of the g_i, Δ is a finite extension field of P in which $\pi(\xi)$ is factorable into linear factors.

As an example, let us take the linear algebra with the defining matrix

$$Q(u) = \begin{bmatrix} e & a \\ a & 5e \end{bmatrix}$$

defined over the rational field P. Then, as in Part 1,

$$Q(\xi) = \begin{bmatrix} x_1 & x_2 \\ x_2 & 5x_1 \end{bmatrix}, \qquad \pi(\xi) = |Q(\xi)| = 5x_1^2 - x_2^2.$$

The parastrophic form is irreducible over the rational field, but over $\Delta=P(5^{1/2})$, $\pi(\xi)=(5^{1/2}x_1-x_2)(5^{1/2}x_1+x_2)$.

As an example of the failure of the theorem for a non-commutative algebra, consider the total matric algebra of order 4 with basis $u=(e_{11}, e_{12}, e_{21}, e_{22})$ over any field P we wish to choose. In this case

([10]) M. Deuring, ibid., p. 32.

([11]) van der Waerden, ibid.

$$Q(u) = \begin{bmatrix} e_{11} & e_{12} & 0 & 0 \\ 0 & 0 & e_{11} & e_{12} \\ e_{21} & e_{22} & 0 & 0 \\ 0 & 0 & e_{21} & e_{22} \end{bmatrix}, \qquad Q(\xi) = \begin{bmatrix} x_1 & x_2 & 0 & 0 \\ 0 & 0 & x_1 & x_2 \\ x_3 & x_4 & 0 & 0 \\ 0 & 0 & x_3 & x_4 \end{bmatrix},$$

$$\pi(\xi) = |Q(\xi)| = (x_2x_3 - x_1x_4)^2.$$

Here $\pi(\xi)$ contains a factor of degree two which is irreducible over all extension fields of P.

Lemma 4.1 may be applied to a theorem of Burnside[12] on the factorization of a certain determinant. Consider the matrix

$$\begin{bmatrix} x_1, & x_2, & \cdots, & x_n \\ x_{\alpha_1}, & x_{\alpha_2}, & \cdots, & x_{\alpha_n} \\ \cdot & \cdot & \cdots & \cdot \\ x_{\sigma_1}, & x_{\sigma_2}, & \cdots, & x_{\sigma_n} \end{bmatrix}$$

where each row is a permutation of the first row, the permutations constituting a commutative group $\mathfrak{G}$ of order n. This matrix is the parastrophic matrix of the group-ring with the elements of $\mathfrak{G}$ in some order as basis elements. Since $\mathfrak{G}$ is commutative, its group-ring is commutative, and the determinant of the parastrophic matrix is factorable into linear factors in some finite extension field of the field of coefficients.

An algebra is called primary if it contains an identity element and its remainder class ring with respect to the radical is simple. A simple commutative algebra is a field; for a left ideal of a commutative algebra is a two-sided ideal. The decomposition of the identity of a commutative algebra into the sum of primitive idempotents induces a decomposition of the algebra into the direct sum of primary ideals. The remainder class ring of each component of this decomposition contains but one idempotent and is therefore a field[13].

We choose to impose upon the field P the condition that it be perfect, i.e., that it admit no inseparable extension.

For convenience we shall define a primary commutative algebra A to be in *normal form* if it is defined over a maximal sub-field as coefficient field.

LEMMA 4.2. *A primary commutative algebra A/P defined over the perfect field P may be put in normal form. We write $A/P = B/F$ where A/P modulo its radical N is isomorphic to the field F.*

The primary commutative algebra A is defined over the perfect field P. We may write $A = F + N$ where N is the radical of A and $A/N \simeq F$[14]. The

[12] W. Burnside, Messenger of Mathematics, vol. 23 (1894), pp. 112–114.

[13] For this paragraph, cf. Deuring, ibid., p. 17; van der Waerden, ibid., p. 47.

[14] M. Deuring, ibid., p. 23.

remainder class ring is a field since A/N is simple. The field F contained in A may be employed as a coefficient field. Let $e_1, e_2, \cdots, e_n$ be a P-basis for A. Consider the set of elements $\phi_1 e_1+\phi_2 e_2+\cdots+\phi_n e_n$ where the ϕ_i range over F. The set of all such numbers is contained in A, and on the other hand, since $F \supseteq P$, it contains A. Then an F-basis may be chosen for A considered as a finite F-module. To indicate this change of basis field, we write

$$A/P = B/F.$$

The field F is maximal in $A=F+N$ since any ring which contains F must also contain elements of the radical N.

LEMMA 4.3. *If $\mathfrak{a}$ is a minimal ideal of the commutative primary algebra A, then $\mathfrak{a}N=0$ where N is the radical of A. If A is written in normal form B/F, $\mathfrak{a}$ is of order 1 over F.*

Let A/P be written in normal form B/F. The same ideal $\mathfrak{a}$ and the same radical N are to be considered since the rings A/P and B/F are identical except for form of expression over different fields.

Let $\mathfrak{a}$ be a minimal ideal of B. Let $b\neq 0$ be an element of $\mathfrak{a}$. Then $\mathfrak{a}=Bb$; for $\mathfrak{a}\supseteq Bb$, and $\mathfrak{a}$ is minimal. Every element of $\mathfrak{a}$ is of the form xb where x lies in B. The correspondence $b\rightarrow cb$, c an element of B, defines an endomorphism([15]) of the ideal $\mathfrak{a}$ considered as an additive group. Let $x_1 b\rightarrow x_1 cb$, $x_2 b\rightarrow x_2 cb$, where x_1, x_2 are in B. Then $x_1 b+x_2 b=(x_1+x_2)b\rightarrow x_1 cb+x_2 cb=(x_1+x_2)cb$. Let $\mathfrak{c}$ be the set of elements of B such that $c\mathfrak{a}=0$. Then $\mathfrak{c}$ is an ideal, and $B/\mathfrak{c}$ is an absolute operator domain for $\mathfrak{a}$. However the endomorphism ring of a minimal ideal is a division algebra([16]), so that the ring $B/\mathfrak{c}$ is a field. Since every ideal of a primary commutative ring is contained in the radical([17]), $\mathfrak{c}\subseteq N$, and $B/\mathfrak{c}\sim B/N\simeq F$. However no field can be homomorphic to a second unless the homomorphism is an isomorphism. Therefore $B/\mathfrak{c}\simeq B/N$, and $\mathfrak{c}=N$, $N\mathfrak{a}=0$. It follows that $\mathfrak{a}$ is of order 1; for

$$(\alpha e+a)b=\alpha eb=\alpha b$$

where α is an element of F, a is an element of N, and $Bb=\mathfrak{a}$, $b\neq 0$ being an element of $\mathfrak{a}$.

THEOREM 4.1. *A necessary and sufficient condition that a commutative primary algebra defined over the perfect field P be a Frobenius algebra is that A have one and only one minimal ideal.*

Suppose that A has one and only one minimal ideal $\mathfrak{a}$. If this ideal is A, then A is a field and the theorem holds. If $\mathfrak{a}$ is a proper ideal, then $\mathfrak{a}$ is of

([15]) van der Waerden, B.L., *Moderne Algebra*, Berlin, Springer, 1936, 2d edition.
([16]) van der Waerden, *Modern Algebra*, vol. 2, p. 165.
([17]) Krull, W., *Idealtheorie*, Berlin, Springer, 1935, p. 22.

order 1 over F by Lemma 4.3 and $\mathfrak{a}$ is of order k over P where k is the order of F over P. Then if $U(\mathfrak{a})$ is defined for A/P and A is of order n over P, $U(\mathfrak{a})$ is of order $n-k$ over P. This follows from Lemma 2.4. Since $\mathfrak{a}$ is minimal, every nonzero ideal of A must contain $\mathfrak{a}$, and therefore every orthogonal module of a nonzero ideal is contained in $U(a)$ (Lemma 2.5). Then if χ is a vector of M not contained in $U(\mathfrak{a})$, $|Q(\chi)| \neq 0$; for by Lemma 3.1 $|Q(\chi)| = 0$ implies that x lies in the orthogonal module of a nonzero ideal in A.

Conversely suppose that A has at least two minimal ideals $\mathfrak{a}$ and $\mathfrak{b}$. Since $\mathfrak{a}$ and $\mathfrak{b}$ contain no ideal other than the zero ideal, they must be principal ideals. Let $\mathfrak{a}=(a)$, $\mathfrak{b}=(b)$ where $a \neq 0$, $b \neq 0$. The elements a and b must lie in the radical of A since every ideal of A lies in its radical.

Let A be written in normal form B/F. Then by Lemma 4.3, $\mathfrak{a}$ and $\mathfrak{b}$ are of order 1 over F. Choose a basis

$$v = (e, a_2, \cdots, a_{m-2}, a, b)$$

where e is the identity of B, and $a_2, \cdots, a_{m-2}$ lie in the radical of B, and m is the order of B over F. By Lemma 4.3 the product of $\mathfrak{a}$ and $\mathfrak{b}$ with any element of the radical is zero. We have

$$Q_B(v) = \begin{bmatrix} e, & a_2, \cdots, a_{m-2}, & a, & b \\ a_2, & a_2^2, \cdots, \cdot, & 0, & 0 \\ \cdot & \cdot \cdots \cdot & \cdot & \cdot \\ a_{m-2}, & \cdot \cdots \cdot, & 0, & 0 \\ a, & 0, \cdots, 0, & 0, & 0 \\ b, & 0, \cdots, 0, & 0, & 0 \end{bmatrix}.$$

Let w be a basis of F with respect to P. Then B/F may be written as A/P with basis

$$u = (ew, a_2w, \cdots, a_{m-2}w, aw, bw),$$

and the corresponding defining matrix takes the form

$$Q_A(u) = \begin{bmatrix} eQ_P(w), & a_2Q_P(w), \cdots, a_{m-2}Q_P(w), & aQ_P(w), & bQ_P(w) \\ a_2Q_P(w), & a_2^2Q_P(w), \cdots, \cdot, & 0, & 0 \\ \cdot & \cdot \cdots \cdot & \cdot & \cdot \\ a_{m-2}Q_P(w), & \cdot \cdots \cdot, & 0, & 0 \\ aQ_P(w), & 0, \cdots, 0, & 0, & 0 \\ bQ_P(w), & 0, \cdots, 0, & 0, & 0 \end{bmatrix}$$

where $Q_P(w)$ is the defining matrix for F with respect to P. The form of $Q_A(u)$

implies immediately that the parastrophic matrix $Q_A(\xi)$ is singular; A cannot be a Frobenius algebra.

COROLLARY 4.1. *If the parastrophic form $\pi(\xi)$ of the commutative primary algebra A contains a linear factor $g(\xi)$, the vector solutions of the equation $g(\xi)=0$ form the orthogonal module of an ideal of order 1 in A.*

Let $\pi(\xi)=g(\xi)h(\xi)$ where $\xi=(x_1, x_2, \cdots, x_n)$ and where $g(\xi)$ is a linear polynomial in $x_1, x_2, \cdots, x_n$. The solutions of $g(\xi)=0$ form a linear sub-space T of the set of vectors M. If M is of order n, T is of order $n-1$.

Suppose that T is not closed on the right by all $S^\tau(x)$. Then $\eta_1 \neq 0$ and $a \neq 0$ exist such that $\eta_1 S^\tau(a)$ is not an element of T, and

$$M = T + P\eta_1 S^\tau(a).$$

If $\eta \in M$, there exist $\eta_2 \in T$ and $\alpha \in P$ such that $\eta = \eta_2 + \alpha\eta_1 S^\tau(a)$. Since η_1 and η_2 are solutions of $g(\xi)=0$ and therefore solutions of $\pi(\xi)=0$ each of these vectors must, by Lemma 3.1 lie in the orthogonal module of some nonzero ideal of A. Let $\eta_1 \in U(\mathfrak{a}_1)$, $\eta_2 \in U(\mathfrak{a}_2)$. Then by Lemma 2.2 $\eta_1 S^\tau(a) \in U(\mathfrak{a}_1)$. The theorem gives the existence of an unique minimal ideal $\mathfrak{b} \neq 0$, and $\mathfrak{b} \subseteq (\mathfrak{a}_1 \wedge \mathfrak{a}_2)$. But $\eta \in U(\mathfrak{a}_1 \wedge \mathfrak{a}_2)$ which is contained in $U(\mathfrak{b})$. It follows that $|Q(\eta)| = 0$, and A is not a Frobenius algebra.

This contradiction gives the closure of T. By Theorem 2.2 T is the orthogonal module of an ideal of order 1 in A.

THEOREM 4.2. *If the commutative algebra A has only principal ideals, A is a Frobenius algebra.*

Suppose that a primary component B of A were not a Frobenius algebra. Then, by Theorem 4.1, B must have at least two minimal ideals. Since a minimal ideal is principal, these must have the form (a) and (b), $a \neq 0$, $b \neq 0$. Every element of (a) is of the form ca where $c \in B$, and every element of (b) is of the form cb.

Consider the ideal (a, b) with elements $c_1a + c_2b$ where $c_i \in B$. Now every element of B can be written as $\phi + d$ where ϕ is an element of the field F, d is an element of the radical N, and $B = F + N$. By assumption the ideal (a, b) is principal and is therefore generated by a fixed element $c_1a + c_2b$, $c_1, c_2 \in B$. By Lemma 4.3 $aN = 0$; and $bN = 0$. It follows that

$$(\phi + d)(c_1a + c_2b) = \phi(c_1a + c_2b).$$

However (a, b) contains the ideals (a) and (b), so that the equations

$$\phi_1(c_1a + c_2b) = a, \quad \phi_2(c_1a + c_2b) = b$$

must have a solution for ϕ_1 and ϕ_2 as elements of F. Since F is a field, this implies that $c_1 = c_2 = 0$ and that $a = b = 0$ contradictory to hypothesis.

Over the perfect field P it is possible to consider the polynomial algebras

of the form

$$A \simeq P[x]/f(x)$$

where $P[x]$ is the polynomial domain in the indeterminant x, and $P[x]/f(x)$ is the remainder class ring with respect to the polynomial $f(x)$. The basis of A may be chosen as $1, x, x^2, \cdots, x^{n-1}$, and reduction modulo $f(x)$ of degree n follows any multiplication. In this case the polynomial ideal theories may be employed([18]). The factorization of $f(x)$ into distinct factors gives a corresponding reduction of A into the direct sum of ideals. If $f(x) = g_1(x)g_2(x)$ where the two factors are relatively prime, then

$$A \simeq P[x]/f(x) \simeq P[x]/g_1(x) + P[x]/g_2(x).$$

If $f(x)$ is the power of an irreducible polynomial, then $P[x]/f(x)$ is primary.

LEMMA 4.4. *If $A \simeq P[x]/f(x)$, A has an ideal of order 1 if and only if $f(x)$ has a linear factor.*

Let $f(x) = g_1(x)g_2(x) \cdots g_r(x)$ where $g_1(x), g_2(x), \cdots, g_r(x)$ are irreducible. Then

$$A = C_1 + C_2 + \cdots + C_r,$$

and $C_i \simeq P[x]/g_i^{\sigma_i}(x)$ is a primary ideal. Now

$$C_i \supseteq F_i \simeq P[x]/g_i(x)$$

so that F_i is a field whose order is the degree of $g_i(x)$. Let $\mathfrak{a}$ be a minimal ideal of C_i in case $\sigma_i > 1$ and C_i itself in case $\sigma_i = 1$. Then $\mathfrak{a}$ is a minimal ideal of A. However, since $\mathfrak{a} \supseteq \mathfrak{a}F_i$, $\mathfrak{a}$ is of order at least that of F_i so that if $\mathfrak{a}$ is of order 1, F_i is of order 1, and $g_i(x)$ is of degree 1.

Conversely if $g_i(x)$ is linear, $C_i \simeq P[x]/g_i^{\sigma_i}(x)$ is isomorphic with $P[y]/y^{\sigma_i}$ where $y = g_i(x)$. The transformation $y = g_i(x)$ has an inverse since $g_i(x)$ is linear. We may choose as a basis of C the elements $1, y, y^2, \cdots, y^{\sigma_i - 1}$. Then, as in the proof of Lemma 4.2, the maximal field contained in C_i is isomorphic with P, and C_i contains an ideal of order 1, or C itself is of order 1.

THEOREM 4.3. *If F is an algebraic extension field of the perfect field P, the parastrophic form of F with respect to its basis over P is irreducible.*

Since P is perfect, $F \simeq P[x]/f(x)$ where $f(x)$ is irreducible over P; for F is generated by a single element satisfying the irreducible polynomial equation $f(x) = 0$. Suppose that for some basis of F

$$\pi(\xi) = g_1(\xi)g_2(\xi)$$

where g_i is of degree $m_i < n$, n being the order of F with respect to P. By Theo-

([18]) Cf. van der Waerden, loc. cit.; Krull, loc. cit.

rem 3.2 $\pi(\xi)$ is a polynomial of degree n in x_1, g_1 is a polynomial of degree m_1 in x_1 over the polynomial field

$$\Gamma = P(x_2, x_3, \cdots, x_n).$$

Let a root μ of g_1 be adjoined to Γ. Then g_1 (and therefore $\pi(\xi)$ has a linear factor in $\Gamma(\mu)$. It follows (Lemma 4.4) that $f(x)$ must have a linear factor over $\Gamma(\mu)$ since $F/\Gamma(\mu)$ must by Corollary 4.1 have an ideal of order 1. However $\pi(\xi)$ is uniquely factorable into linear factors over a finite extension field Δ of P (Lemma 4.1), so that

$$\Gamma(\mu) = P(\theta)(x_2, x_3, \cdots, x_n)$$

where θ is algebraic of degree m_1 with respect to P. Then $f(x)$ has a zero in $P(\theta)$. However $f(x)$ is of degree n over P while m_1 was assumed to be less than n. This contradiction gives the irreducibility of $\pi(\xi)$.

LEMMA 4.5. *If the primary commutative Frobenius algebra A is defined over the perfect field P, and if the maximal field F contained in A is of degree k over P, then A/Ω, where Ω is the algebraic closure of P, has exactly k minimal ideals.*

If F is the maximal field contained in A, then $F \simeq A/N$ where N is the radical of A, and we may write A as the direct sum of F and N([19]);

$$A/P = F/P + N/P$$

where the direct sum is in the sense of P-modules.

Consider that P is perfect and that F is a finite extension field of P. It follows that F is obtained from P by the adjunction of a single element and that this element satisfies an irreducible algebraic equation $f(x) = 0$ of degree k with coefficients in P. Then

$$F/P \simeq P[x]/f(x).$$

If we extend P to its algebraic closure Ω, then

$$A/\Omega = F/\Omega + N/\Omega$$

and, furthermore,

$$F/\Omega \simeq \Omega[x]/f(x).$$

Over Ω, however, $f(x)$ factors into linear factors, and F/Ω decomposes into the direct sum of primary ideals, each containing an unique idempotent. There can be but k idempotents in A/Ω, and each induces a primary component of A/Ω.

Since A/P is a Frobenius algebra, A/Ω is a Frobenius algebra. Each primary component of A/Ω is a Frobenius algebra, and each, by Theorem 4.1,

([19]) M. Deuring, loc. cit., p. 23.

can contain one and only one minimal ideal. We have proved that the number of these components is k since each contains a primitive idempotent of F/Ω. Therefore there are just k distinct minimal ideals in A/Ω.

THEOREM 4.4. *If the primary commutative algebra A is a Frobenius algebra, the parastrophic form is $\pi(\xi)=g^m(\xi)$ where $g(\xi)$ may be chosen as the parastrophic form of $F\simeq A/N$ where N is the radical of A.*

Let A/P be written in normal form B/F with basis $v=(b_1=e, b_2, \cdots, b_m=b)$ where e is the identity element and b is the basis element with respect to F of the unique minimal ideal of B (cf. Lemmas 4.1, 4.3) and $b_2, \cdots, b_m$ lie in the radical of B. The defining matrix of B/F under these conditions is

$$Q_B(v) = \begin{bmatrix} e, & b_2, & \cdots, & b \\ b_2, & b_2^2, & \cdots, & 0 \\ \cdot & \cdot & \cdots & \cdot \\ b, & 0, & \cdots, & 0 \end{bmatrix}.$$

The product of b with any element of the radical is zero by Lemma 4.3. Column m and row m of $Q_B(v)$ consist of zeros except in the first place of each.

Let w be a basis of F with respect to P. Then a basis of $A/P=B/F$ with respect to P will be given by

$$u = (ew, b_2w, \cdots, bw),$$

and the defining matrix of A/P on the basis u is

$$Q_A(u) = \begin{bmatrix} eQ_F(w), & b_2Q_F(w), & \cdots, & bQ_F(w) \\ b_2Q_F(w), & b_2^2Q_F(w), & \cdots, & 0 \\ \cdot & \cdot & \cdots & \cdot \\ bQ_F(w), & 0, & \cdots, & 0 \end{bmatrix}$$

where $Q_F(w)$ is the defining matrix of F with respect to P.

To form the parastrophic matrix $Q_A(\xi)$ from $Q_A(u)$ we substitute $\xi=(x_1, x_2, \cdots, x_n)$ for u in $Q_A(u)$. This substitution will also take place in the upper right-hand corner of the matrix, namely within the matrix $bQ_F(w)=Q_F(bw)$. Since b is linearly independent of the elements of the basis w, $Q_F(bw)$ will be the same except for notation as $Q_F(w)$. Substitution of variables $x_1, x_2, \cdots, x_n$ in $Q_F(bw)$ will, therefore, give the same result as substitution in $Q_F(w)$.

Then $|Q_F(\xi)|=g(\xi)$ must be the parastrophic form of F for the basis w. In the expansion of $|Q_A(\xi)|$ by minors as indicated by the matrix $Q_A(u)$, $g(\xi)$ must enter as a factor. The parastrophic form of A is $\pi(\xi)=g(\xi)f(\xi)$ where $g(\xi)$ is, by Theorem 4.3, irreducible.

We now extend the basis field P to its algebraic closure Ω. Over Ω, how-

ever, $g(\xi)=h_1(\xi)h_2(\xi)\cdots h_k(\xi)$ where k is the order of F over P and the $h_i(\xi)$ are distinct linear polynomials in $x_1, x_2, \cdots, x_n$ (cf. Lemma 4.1). By Lemma 4.5, A/Ω has exactly k minimal ideals. Suppose that $f(\xi)$ contained a linear factor distinct from the $h_i(\xi)$. Then by Corollary 4.1 this factor would give rise to a $(k+1)$st minimal ideal distinct from the rest. Since this is impossible, $f(\xi)$ must contain only the $h_i(\xi)$ as factors. Since $g(\xi)$ is irreducible over P, $f(\xi)$ must be a power of $g(\xi)$, and $\pi(\xi)=g^m(\xi)$ where m is the order of B over F.

5. **Factorization of the rational prime number p.** Let H be a finite extension field of degree n over the rational field R. Let G be the ring of rational integers, and let K be a domain of integrity in H with G-basis $e_1, e_2, \cdots, e_n$. If p is a rational prime number, $\mathfrak{a}=Kp$ is the corresponding principal ideal in K. Then

$$\mathfrak{a} = \mathfrak{q}_1\mathfrak{q}_2\cdots\mathfrak{q}_s$$

where $\mathfrak{q}_i$ is a primary ideal belonging to the prime ideal $\mathfrak{p}_i$, and

$$C = K/\mathfrak{a} \simeq K/\mathfrak{q}_1 + K/\mathfrak{q}_2 + \cdots + K/\mathfrak{q}_s$$

is the direct sum of the primary ideals $K/\mathfrak{q}_i$[20]. This ring is composed of linear combinations of the elements e_i with coefficients in G taken modulo p, and is therefore a linear algebra over the field $P=G/p$, which is a perfect field[21]. Furthermore, C has the same multiplication table as K. If $\pi(\xi)$ is the parastrophic form in the basis $e_1, e_2, \cdots, e_n$,

$$\pi^*(\xi) = \pi(\xi) \pmod{p}$$

is that of the algebra C.

The additive components of C induce a factorization of $\pi^*(\xi)$ into distinct factors. Each component is primary, and the corresponding factor of the parastrophic form is the power of an irreducible factor or is identically zero (Theorem 4.4), i.e., if $\pi(\xi)\not\equiv 0 \pmod{p}$, then

$$\pi(\xi) = g_1^{\sigma_1}g_2^{\sigma_2}\cdots g_s^{\sigma_s} \pmod{p}$$

where the degree of g_i is the degree of the field $K/\mathfrak{p}_i$ over P, and the radical of $K/\mathfrak{q}_i$ is of index at most σ_i.

Two ideals are said to be divisor prime in K if their greatest common divisor is the unit ideal.

LEMMA 5.1. *If the ideal $\mathfrak{a}$ is divisor prime to the conductor $\mathfrak{f}$ of K, then $K/\mathfrak{a}$ is a principal ideal ring.*

By Krull[22] if $\mathfrak{a}$ is divisor prime to $\mathfrak{f}$, $\mathfrak{a}=\mathfrak{p}_1^{\sigma_1}\mathfrak{p}_2^{\sigma_2}\cdots\mathfrak{p}_s^{\sigma_s}$. Let $\mathfrak{b}\supset\mathfrak{a}$. Then

[20] Cf. van der Waerden, loc. cit., vol. 2, pp. 47–50.
[21] van der Waerden, loc. cit., vol. 1, first edition, p. 118.
[22] Krull, ibid.

$(\mathfrak{b}, \mathfrak{f}) \supseteq (\mathfrak{a}, \mathfrak{f}) \supseteq K$, and every ideal which contains $\mathfrak{a}$ is divisor prime to $\mathfrak{f}$. Then $\mathfrak{b} = \mathfrak{p}_1^{\beta_1}\mathfrak{p}_2^{\beta_2} \cdots \mathfrak{p}_s^{\beta_s}$ where $\beta_i \leqq \alpha_i$.

Consider that the ideals of $K/\mathfrak{a}$ are the rings $\mathfrak{b}/\mathfrak{a}$ where $\mathfrak{b} \supset \mathfrak{a}$. Then if $\mathfrak{c} = \mathfrak{p}_1^{\beta_1+1}\mathfrak{p}_2^{\beta_2+1} \cdots \mathfrak{p}_s^{\beta_s+1}$, $\mathfrak{c}_i = \mathfrak{p}_1^{\beta_1+1}\mathfrak{p}_2^{\beta_2+1} \cdots \mathfrak{p}_i^{\beta_i} \cdots \mathfrak{p}_s^{\beta_s+1}$, we may choose d_i in $\mathfrak{c}_i$ and not in $\mathfrak{c}$. Then d_i is contained in every $\mathfrak{p}_j^{\beta_j+1}$ except $\mathfrak{p}_i^{\beta_i+1}$. Furthermore d_i is contained in all $\mathfrak{p}_j^{\beta_j}$. Therefore d_i and $d = d_1 + d_2 + \cdots + d_s$ are contained in $\mathfrak{b}$. It follows that the ideal $(\mathfrak{a}, d)$ which contains $\mathfrak{a}$ is contained in $\mathfrak{b}$. Since $(\mathfrak{a}, d)$ is a divisor of $\mathfrak{a}$, $(\mathfrak{a}, d) = \mathfrak{p}_1^{\gamma_1}\mathfrak{p}_2^{\gamma_2} \cdots \mathfrak{p}_s^{\gamma_s}$, where $\gamma_i \geqq \beta_i$. Then $d \in \mathfrak{p}^{\gamma_i}$. However d is not divisible by $\mathfrak{p}_i^{\beta_i+1}$ so that $\gamma_i \leqq \beta_i$. It follows that $\gamma_i = \beta_i$, and

$$(\mathfrak{a}, d) = \mathfrak{p}_1^{\beta_1}\mathfrak{p}_2^{\beta_2} \cdots \mathfrak{p}_s^{\beta_s} = \mathfrak{b}.$$

The ideal $\mathfrak{b}/\mathfrak{a}$ of $K/\mathfrak{a}$ is generated by the element d, i.e., $K/\mathfrak{a}$ is a principal ideal ring(23).

When K is the maximal domain of integrity of H, the conductor $\mathfrak{f}$ is the unit ideal, and $K/\mathfrak{a}$ is principal for every ideal in K.

THEOREM 5.1. *If* $\pi(\xi) \equiv g_1^{\sigma_1}g_2^{\sigma_2} \cdots g_s^{\sigma_s} \pmod{p}$, *then*

$$\mathfrak{a} = Kp = \mathfrak{q}_1\mathfrak{q}_2 \cdots \mathfrak{q}_s$$

where the $\mathfrak{q}_i$ *are primary ideals belonging to the prime ideals* $\mathfrak{p}_i$ *of degree that of* g_i. *If* $\pi(\xi) \equiv 0 \pmod{p}$, p *is not prime to the conductor of* K.

The theorem follows immediately from the above conclusions and from Theorem 4.2.

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

(23) The proof of Lemma 5.1 is taken from van der Waerden, loc. cit.

ON A THEOREM OF SCHUR AND ON FRACTIONAL INTEGRALS OF PURELY IMAGINARY ORDER

BY

H. KOBER

1. Let $L_p(a, b)$ be the space of all functions $f(y)$ whose pth power is integrable over (a, b) or which are measurable and essentially bounded over (a, b) for $1 \leqq p < \infty$ or $p = \infty$ respectively, with the norm

$$|f|_p = \left\{ \int_a^b |f(y)|^p dy \right\}^{1/p} \qquad [1 \leqq p < \infty],$$

$$|f|_p = \operatorname*{ess.\,u.b.}_{a<y<b} |f(y)| \qquad [p = \infty];$$

let $p' = p/(p-1)$ and

$$L_p = L_p(0, \infty).$$

The following theorem is in substance due to I. Schur(1):

Let $K(x, y)$ be homogeneous of degree -1 and $K(x, y) \geqq 0$ for $0 < x < \infty$, $0 < y < \infty$, let $K(x, y)x^{-1/2} \subset L_1$, and let $f(x) \subset L_2$; then

$$\left| \int_0^\infty K(x, y) f(x) dx \right|_2 \leqq \kappa |f(y)|_2,$$

where

$$\kappa = \int_0^\infty K(x, 1) x^{-1/2} dx = \int_0^\infty K(1, y) y^{-1/2} dx.$$

The constant κ is the best possible.

Of course the inequality is true when $K(x, y)$ takes negative or even complex values also, if we replace κ by

$$\tilde{\kappa} = \int_0^\infty |K(x, 1)| x^{-1/2} dx = \int_0^\infty |K(1, y)| y^{-1/2} dy.$$

However $\tilde{\kappa}$ is not the best possible constant any more. We shall give a better

Presented to the Society, February 22, 1941; received by the editors May 13, 1940; §§8–11 added June 21, 1940.

(1) Journal für die reine und angewandte Mathematik, vol. 140 (1911), pp. 1–28. The corresponding theorem for $f \subset L_p$, $1 < p < \infty$, was proved by G. H. Hardy, J. E. Littlewood, and G. Pólya, vide *Inequalities*, Cambridge, 1934, Theorem 319.

theorem for this case and we shall use it to deal with fractional integrals the order of which is an imaginary number, thus filling a gap in the literature.

Throughout this paper we denote constants depending on the given parameters by the single symbol C; α and β denote finite numbers such that $\Re(\alpha)>0$, $\Re(\beta)=0$.

2. THEOREM I. *Let* (i) $K(x, y)$ *be homogeneous of degree* -1, (ii) $K(x, 1)x^{-1/2} \subset L_1$, (iii) $f(x) \subset L_2$; *then the function*

$$Wf = \int_0^\infty K(x, y)f(x)dx$$

exists for almost all values of y *in* $(0, \infty)$, *and*

$$|Wf|_2 \leqq \kappa_0 |f(y)|_2 = \max_{-\infty<\tau<\infty} |\omega(\tau)| \cdot |f(y)|_2,$$

where

$$\omega(\tau) = \int_0^\infty K(x, 1)x^{-1/2-i\tau}dx = \int_0^\infty K(1, y)y^{-1/2+i\tau}dy, \qquad \kappa_0 = \max_{-\infty<\tau<\infty} |\omega(\tau)|.$$

The constant is the best possible.

Obviously $\kappa_0 \leqq \tilde{\kappa}$; when $K(x, y) \geqq 0$ then $\kappa_0 = \kappa$, as we may see taking $\tau = 0$.

Proof. Without loss of generality we may suppose $K(x, y)$ to be no null-function; then $\kappa_0 > 0$. Let $1 < a < \infty$, $f(x, a) = f(x)$ in (a^{-1}, a), $f(x, a) = 0$ otherwise, and let

$$M\phi = \text{l.i.m. sq.}_{N\to\infty} \int_{1/N}^N \phi(x)x^{-1/2+i\tau}dx \qquad [\phi \subset L_2].$$

In consequence of Schur's theorem W is a bounded linear transformation in L_2; the Mellin transform M is a bounded linear transformation from $L_2(0, \infty)$ into $L_2(-\infty, \infty)$. We have

$$\begin{aligned}\int_0^\infty y^{-1/2+i\tau}W\{f(x, a)\}dy &= \int_0^\infty y^{-1/2+i\tau}dy \int_{1/a}^a K(x, y)f(x)dx \\ &= \int_{1/a}^a f(x)x^{-1/2+i\tau}dx \int_0^\infty K(1, v)v^{-1/2+i\tau}dv\end{aligned} \tag{2.1}$$

when we put $y = vx$ and make use of the homogeneousness of $K(x, y)$; the interchanging of the integrations is justified by absolute convergence of the right-hand repeated integral. Since the left-hand integral exists, it must be equal to $MW\{f(x, a)\}$, therefore we have

$$MW\{f(x, a)\} = \omega(\tau)M\{f(x, a)\}.$$

Since $|f(x, a)-f(x)|_2 \to 0$ $[a \to \infty]$, and $|\omega(\tau)| \leqq \kappa_0 < \infty$, by the continuity of the operations M and W we get

$$MWf = \omega(\tau)Mf; \tag{2.2}$$

therefore $|MWf|_2 \leqq \kappa_0 |Mf|_2$ in $L_2(-\infty, \infty)$. Now the operator $(2\pi)^{-1/2}M$ is isometric, and so we obtain the first assertion of the theorem.

The function $\omega(\tau)$ is continuous in consequence of (ii) and attains its maximum value at a finite point τ, since, by the Riemann-Lebesgue theorem, $\omega(\tau) \to 0$ $[\tau \to \pm \infty]$. Now let λ be any positive number smaller than κ_0. Then we can easily show the existence of functions $f(x) \subset L_2$ such that $|Wf|_2 > \lambda |f|_2$.

Let E be a set of measure $m(E) > 0$ such that $|\omega(\tau)| > \lambda$ in E and E is included in some finite interval. Take $\phi(\tau) = 1$ in E and $\phi(\tau) = 0$ otherwise, and let $f = M^{-1}\phi$. Then from (2.2) we have

$$\int_{-\infty}^{\infty} |MWf|^2 d\tau = \int_E |\omega(\tau)|^2 d\tau > \lambda^2 m(E)$$

$$= \lambda^2 \int_{-\infty}^{\infty} |\phi(\tau)|^2 d\tau = 2\pi\lambda^2 \int_0^{\infty} |f(x)|^2 dx,$$

$$\int_0^{\infty} |Wf|^2 dy > \lambda^2 \int_0^{\infty} |f(x)|^2 dx.$$

Hence the theorem is proved.

3. We could give an alternative proof by the theory of "general transforms," without making use of Schur's theorem. Let V be a transformation of the form

$$Vf = \int_0^{\infty} L(x, y) f(x) dx,$$

the infinite integral being defined in some sense. Then it turns out that, roughly speaking, the class of all transformations which are representable in the form $V_1 V_2$ is identical with the class of the transformations

$$Wf = \int_0^{\infty} K(x, y) f(x) dx,$$

where $K(x, y)$ is homogeneous of degree -1. We leave that proof of I to the reader(2).

(2) W belongs to the so-called "product-class." We need the lemmas:

A. *Let $y^{-1}\chi(y) \subset L_2$, let $\omega(\tau) = (\frac{1}{2} - i\tau) M\{y^{-1}\chi(y)\}$ be essentially bounded in $(-\infty, \infty)$, and let $\chi(y)$ have the form $\chi(y) = \int_1^y H(\xi) d\xi + c$, where $c = \chi(1)$ is an arbitrary constant. Then, for any $f \subset L_2$, the function*

$$g(y) = Wf = \operatorname{l.i.m.}_{N \to \infty} \text{sq.} \int_{1/N}^{N} H\left(\frac{y}{x}\right) f(x) \frac{dx}{x}$$

4. We replace the customary operators(3)

$$f_\alpha^+(y) = \frac{1}{\Gamma(\alpha)} \int_0^y (y-x)^{\alpha-1} f(x)dx, \tag{4.1}$$

$$f_\alpha^-(y) = \frac{1}{\Gamma(\alpha)} \int_y^\infty (x-y)^{\alpha-1} f(x)dx \tag{4.2}$$

by the more general ones

$$f_{\eta,\alpha}^+(y) = I_{\eta,\alpha}^+ f = \frac{y^{-\eta-\alpha}}{\Gamma(\alpha)} \int_0^y (y-x)^{\alpha-1} x^\eta f(x)dx, \tag{4.3}$$

$$f_{\eta,\alpha}^-(y) = J_{\eta,\alpha}^- f = \frac{y^\eta}{\Gamma(\alpha)} \int_y^\infty (x-y)^{\alpha-1} x^{-\eta-\alpha} f(x)dx, \tag{4.4}$$

where η is a given parameter. Obviously

$$f_\alpha^+(y) = y^\alpha I_{0,\alpha}^+ f, \qquad f_\alpha^-(y) = y^\alpha J_{-\alpha,\alpha}^- f. \tag{4.5}$$

In another paper we have proved that $I_{\eta,\alpha}^+ f$ and $J_{\eta,\alpha}^- f$ are bounded linear transformations in L_p for $1 \leqq p \leqq \infty$ when $\Re(\alpha) > 0$ and when $\Re(\eta) > -1/p'$ or $\Re(\eta) > -1/p$ respectively(4). Obviously the definitions above have no meaning at all when we replace α by an imaginary number β, but we shall show that the operators $I_{\eta,\beta}^+ f$ and $J_{\eta,\beta}^- f$ exist in some sense for any $f \subset L_2$. Those definitions are of importance in the theory of Hankel transforms, as will be shown in a joint paper of A. Erdélyi and myself.

---

exists, and $Mg = \omega(\tau)Mf$.

Vide H. Kober, Quarterly Journal of Mathematics, Oxford, vol. 8 (1937), pp. 172–185, §6 and Theorem 2A.

B. *Let* $\phi(x) \subset L_1$ *and* $\psi(y) = y^{-1}\int_0^y \phi(x)x^{1/2}dx$; *then* $y^{1/2}\psi(y) \to 0$ *for* $y \to 0$ *and* $y \to \infty$, *and* $|\psi|_2 \leqq |\phi|_1$.

C. *Let* $y^{-1}\chi(y) \subset L_2$, *let* $y^{-1/2}\chi(y) \to 0$ *for* $y \to 0$ *and* $y \to \infty$, *and let* $\chi(y)$ *have the form as in Lemma* A; *then*

$$(\tfrac{1}{2} - ir)M\{y^{-1}\chi(y)\} = \lim_{N\to\infty} \int_{1/N}^N H(y)y^{-1/2+ir}dy,$$

if the right-hand limit exists.

Cf. H. Kober, loc. cit., Theorem 3(i).

(3) Cf. H. Weyl, Vierteljahrsschrift der Naturforschende Gesellschaft, Zurich, vol. 62 (1917), pp. 296–302; G. H. Hardy and J. E. Littlewood, Mathematische Zeitschrift, vol. 27 (1928), pp. 565–606; E. R. Love and L. C. Young, Proceedings of the London Mathematical Society, (2), vol. 44 (1938), pp. 1–28. The operator $y^{-\alpha} f_\alpha^-(y)$ exists in L_p with domain L_p and is bounded when $1 \leqq p < \{\Re(\alpha)\}^{-1}$. Also cf. J. D. Tamarkin, Annals of Mathematics, (2), vol. 31 (1930), pp. 219–228.

(4) Cf. *Inequalities*, Theorem 329. H. Kober, Quarterly Journal of Mathematics, Oxford, vol. 11 (1940), pp. 193–211.

Let $\Re(\eta) > -1/2$ and(4)

$$K(x, y) = \begin{cases} \{\Gamma(\alpha)\}^{-1}(y - x)^{\alpha-1}x^{\eta}y^{-\eta-\alpha} & [0 < x < y], \\ 0 & [x > y]; \end{cases}$$

then $K(x, y)$ satisfies the hypotheses of Theorem I, and we have

$$\omega(\tau) = \int_0^\infty K(x, 1)x^{-1/2-i\tau}dx = \frac{1}{\Gamma(\alpha)}\int_0^1 (1 - x)^{\alpha-1}x^{\eta-1/2-i\tau}dx$$

$$= \frac{\Gamma(\eta + \frac{1}{2} - i\tau)}{\Gamma(\eta + \alpha + \frac{1}{2} - i\tau)},$$

$$I^+_{\eta,\alpha}f = \int_0^\infty K(x, y)f(x)dx \qquad [f \subset L_2],$$

therefore by the theorem

$$|I^+_{\eta,\alpha}f|_2 \leqq \max_{-\infty<\tau<\infty} \left|\frac{\Gamma(\eta + \frac{1}{2} - i\tau)}{\Gamma(\eta + \alpha + \frac{1}{2} - i\tau)}\right| |f|_2 = \kappa_0 |f|_2.$$

When we take $|\alpha| < C$, then, in consequence of a well known property of the gamma function, κ_0 is uniformly bounded for $\Re(\alpha) > 0$; therefore

$$|I^+_{\eta,\alpha}f|_2 \leqq C|f|_2,$$

where C depends on η only. Let β be any fixed imaginary number or zero; then, by a well known theorem on weak convergence, a sequence $\alpha_1, \alpha_2, \alpha_3, \cdots$ and a function $\phi(y) \subset L_2$ exist such that $I^+_{\eta,\alpha_n}f$ converges weakly to $\phi(y)$ when α_n tends to β $[n \to \infty]$.

A similar argument applies to $J^-_{\eta,\alpha}f$, and we now define

$$I^+_{\eta,\beta}f = \text{weak limit}_{\alpha_n \to \beta} I^+_{\eta,\alpha_n}f \qquad [\Re(\eta) > -\tfrac{1}{2}, f \subset L_2] \tag{4.6}$$

$$J^-_{\eta,\beta}f = \text{weak limit}_{\alpha_m \to \beta} J^-_{\eta,\alpha_m}f \tag{4.7}$$

for some sequences $\{\alpha_n\}$, $\{\alpha_m\}$.

5. **Strong convergence.** Starting from $I^+_{\eta,\beta}\psi$ and $J^-_{\eta,\beta}\psi$ for step-functions ψ we can show that $I^+_{\eta,\alpha}f$ or $J^-_{\eta,\alpha}f$ converges to $I^+_{\eta,\beta}f$ or $J^-_{\eta,\beta}f$ in the strong sense also for any $f \subset L_2$ when α tends to β. We can also proceed in a shorter way. By (2.2) we have

$$MI^+_{\eta,\alpha}f = \Gamma(\eta + \tfrac{1}{2} - i\tau)\{\Gamma(\eta + \alpha + \tfrac{1}{2} - i\tau)\}^{-1}Mf \tag{5.1}$$

and, taking

$$K(x, y) = \begin{cases} 0 & [0 < x < y] \\ \{\Gamma(\alpha)\}^{-1}(x-y)^{\alpha-1}x^{-\eta-\alpha}y^{\eta} & [x > y], \end{cases}$$

we get

$$MJ^{-}_{\eta,\alpha}f = \Gamma(\eta + \tfrac{1}{2} + i\tau)\{\Gamma(\eta + \alpha + \tfrac{1}{2} + i\tau)\}^{-1}Mf. \tag{5.2}$$

Let

$$Mf = g(\tau);\ \omega(\tau;\alpha) = \Gamma(\eta + \tfrac{1}{2} - i\tau)/\Gamma(\eta + \alpha + \tfrac{1}{2} - i\tau).$$

Then

$$\int_{-\infty}^{\infty} |\, MI^{+}_{\eta,\alpha}f - \omega(\tau;\beta)g(\tau)\,|^2 d\tau = \int_{-\infty}^{\infty} |\, g(\tau)\,|^2 |\, \omega(\tau;\alpha) - \omega(\tau;\beta)\,|^2 d\tau$$

$$= \int_{-\infty}^{-N} + \int_{N}^{\infty} + \int_{-N}^{N} = Z_1 + Z_2 + Z_3.$$

Since $\omega(\tau;\alpha)$ is bounded in $(-\infty, \infty)$ uniformly when $\Re(\alpha) > 0$ and $|\alpha| < C$, and since $g(\tau) \subset L_2(-\infty, \infty)$, we can fix N sufficiently large such that $Z_1 < \epsilon/3$, $Z_2 < \epsilon/3$ uniformly in α for any given $\epsilon > 0$. Now it is easy to show that $Z_3 < \epsilon/3$ when $|\alpha - \beta|$ is sufficiently small. Hence $MI^{+}_{\eta,\alpha}f$ converges strongly to the function $\omega(\tau;\beta)g(\tau)$ and, by the property of the Mellin transformation mentioned above, $I^{+}_{\eta,\alpha}f$ to $M^{-1}\{\omega(\tau;\beta)g(\tau)\}$. By the same argument we get the corresponding result for $J^{-}_{\eta,\alpha}f$, and so we have

THEOREM II. *Let $\Re(\eta) > -1/2$, let $\Re(\alpha) > 0$ and $\Re(\beta) = 0$ and $\alpha \to \beta$, and let $f \subset L_2$; then the functions $I^{+}_{\eta,\alpha}f$ and $J^{-}_{\eta,\alpha}f$ converge strongly to $I^{+}_{\eta,\beta}f$ and $J^{-}_{\eta,\beta}f$ respectively, where*

$$I^{+}_{\eta,\beta}f = M^{-1}\left\{\frac{\Gamma(\eta + \tfrac{1}{2} - i\tau)}{\Gamma(\eta + \beta + \tfrac{1}{2} - i\tau)}Mf\right\}, \tag{5.3}$$

$$J^{-}_{\eta,\beta}f = M^{-1}\left\{\frac{\Gamma(\eta + \tfrac{1}{2} + i\tau)}{\Gamma(\eta + \beta + \tfrac{1}{2} + i\tau)}Mf\right\}. \tag{5.4}$$

Evidently $I^{+}_{\eta,0}f = f$, $J^{-}_{\eta,0}f = f$.

6. **The inversions of the operators $I^{+}_{\eta,\beta}f$, $J^{-}_{\eta,\beta}f$.** The operators $(I^{+}_{\eta,\beta})^{-1}$ and $(J^{-}_{\eta,\beta})^{-1}$ are also bounded linear transformations in L_2. We have

THEOREM III. *Let $\Re(\eta) > -1/2$, $\Re(\beta) = 0$, let $f(x) \subset L_2$, and let*

$$I^{+}_{\eta,\beta}f = g(y); \qquad J^{-}_{\eta,\beta}f = h(y). \tag{6.1}$$

Then

$$f = I^{+}_{\eta+\beta,-\beta}g; \qquad f = J^{-}_{\eta+\beta,-\beta}h. \tag{6.2}$$

The proof follows from (5.3) and (5.4), immediately; for instance,

$$MI^{+}_{\eta+\beta,-\beta}g = \frac{\Gamma(\eta+\beta+\frac{1}{2}-i\tau)}{\Gamma(\eta+\beta-\beta+\frac{1}{2}-i\tau)}Mg = Mf.$$

From (5.3) and (5.4) we may also see that both the domain and the range of $(I^{+}_{\eta,\beta})^{-1}$ and $(J^{-}_{\eta,\beta})^{-1}$ are L_2, since $\{\omega(\tau;\beta)\}^{-1}$ is bounded in $(-\infty, \infty)$.

Of course the operators $I^{+}_{\eta,\alpha}$ and $J^{-}_{\eta,\alpha}$ do not possess this simply property.

7. **Application to the customary fractional integrals, to that of Riemann-Liouville and to that of Weyl.** Let the operators $X_\alpha f = f^{+}_{\alpha}(y)$ and $Y_\alpha f = f^{-}_{\alpha}(y)$ be defined by

$$X_\beta f = y^{\beta} I^{+}_{0,\beta} f, \qquad Y_\beta f = y^{\beta} J^{-}_{-\beta,\beta} f \qquad [f \subset L_2] \tag{7.1}$$

when they are of imaginary order, in accordance with (4.5). Since $|y^{\beta}| = 1$, X_β and Y_β are bounded linear transformations in L_2 with domain L_2, and it is not difficult to show that, for $\alpha \to \beta$,

$$|y^{\beta-\alpha} f^{+}_{\alpha} - X_\beta|_2 \to 0, \qquad |y^{\beta-\alpha} f^{-}_{\alpha} - Y_\beta|_2 \to 0.$$

The semi-group property of f^{+}_{α} in $L_p(0, a)$ for $1 \leqq p \leqq \infty$, $0 < a < \infty$ is well known(8). Here we shall prove

THEOREM IV. *The transformations X_β or Y_β form a group in L_2.*

Since $X_0 f = I^{+}_{0,0} f = f$ and $Y_0 f = J^{-}_{0,0} f = f$, we have only to prove that, for any imaginary numbers β, γ

$$X_\beta X_\gamma = X_{\beta+\gamma}; \qquad Y_\beta Y_\gamma = Y_{\beta+\gamma}. \tag{7.2}$$

We need the following lemmas:

LEMMA 1. *When $f \subset L_2$, $\Re(\eta) > -1/2$, $\Re(\lambda) \geqq 0$, $\Re(\mu) \geqq 0$,*

$$I^{+}_{\eta+\lambda,\mu} I^{+}_{\eta,\lambda} = I^{+}_{\eta,\lambda+\mu} = I^{+}_{\eta,\lambda} I^{+}_{\eta+\lambda,\mu}; \qquad J^{-}_{\eta+\lambda,\mu} J^{-}_{\eta,\lambda} = J^{-}_{\eta,\lambda+\mu} = J^{-}_{\eta,\lambda} J^{-}_{\eta+\lambda,\mu}. \tag{7.3}$$

LEMMA 2. *When $f \subset L_2$, $\Re(\lambda) \geqq 0$, $\Re(\nu) = 0$,*

$$I^{+}_{\eta,\lambda} f = y^{\nu} I^{+}_{\eta+\nu,\lambda}\{x^{-\nu} f(x)\}; \qquad J^{-}_{\eta,\lambda} f = y^{-\nu} J^{-}_{\eta+\nu,\lambda}\{x^{\nu} f(x)\}. \tag{7.4}$$

We can easily prove Lemma 1 by taking the Mellin transforms of both sides and employing (5.1)–(5.4).

The proof of Lemma 2 follows from the definitions (4.3) and (4.4) immediately when $\Re(\lambda) > 0$, since $x^{\nu} f(x) \subset L_2$, $x^{-\nu} f(x) \subset L_2$. Taking $\Re(\lambda) > 0$, $\lambda \to \beta$, we have

(8) E. Hille, Annals of Mathematics, (2), vol. 40 (1939), 4.4. In this paper the theory of semi-groups is developed. Cf. E. Hille, Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 159–161.

$$|I^+_{\eta,\lambda} - I^+_{\eta,\beta}|_2 \to 0;$$

$$|y^\nu I^+_{\eta+\nu,\lambda}\{x^{-\nu}f(x)\} - y^\nu I^+_{\eta+\nu,\beta}\{x^{-\nu}f(x)\}|_2 = |I^+_{\eta+\nu,\lambda}\{x^{-\nu}f(x)\} - I^+_{\eta+\nu,\beta}\{x^{-\nu}f(x)\}|_2 \to 0,$$

and so (7.3) is true for $\lambda=\beta$ also.

Now by (7.1)

$$X_\beta X_\gamma f = y^\beta I^+_{0,\beta}\{x^\gamma I^+_{0,\gamma}f\},$$

and $I^+_{0,\beta}\{x^\gamma\phi\} = y^\gamma I^+_{\gamma,\beta}\phi$ by Lemma 2. Hence, by Lemma 1,

$$X_\beta X_\gamma f = y^{\beta+\gamma} I^+_{\gamma,\beta} I^+_{0,\gamma} f = y^{\beta+\gamma} I^+_{0,\beta+\gamma} f = X_{\beta+\gamma} f.$$

Similarly we have

$$Y_\beta Y_\gamma f = y^\beta J^-_{-\beta,\beta}\{x^\gamma J^-_{-\gamma,\gamma}f\} = y^{\beta+\gamma} J^-_{-\beta-\gamma,\beta} J^-_{-\gamma,\gamma} f = y^{\beta+\gamma} J^-_{-\beta-\gamma,\beta+\gamma} f = Y_{\beta+\gamma} f.$$

COROLLARY. *The transformations $(X_\beta)^{-1}$ and $(Y_\beta)^{-1}$ are linear and bounded in L_2 with domain L_2, and $(X_\beta)^{-1} = X_{-\beta}$, $(Y_\beta)^{-1} = Y_{-\beta}$.*

8. We shall now deal with the corresponding problems in L_p for $p \geqq 1$. We do not know if Theorem I can be generalized in some way for $p \neq 2$. Therefore we cannot extend the results of §§4–7 to the general case $f \subset L_p$ $[1 \leqq p \leqq \infty]$. We have to restrict ourselves to certain subspaces of L_p or, as in Theorem VI, to the case when α tends to zero under certain conditions. Moreover we shall discuss the characteristic values (§10).

Let $0 < a < \infty$ and let the step-function $\phi_a(x)$ be defined by $\phi_a(x) = 1$ for $0 \leqq x \leqq a$, $\phi_a(x) = 0$ otherwise. When $\Re(\zeta) > -1$, we easily find

$$(8.11) \qquad I^+_{\xi,\alpha}\phi_a = \frac{\Gamma(\xi+1)}{\Gamma(\xi+\alpha+1)} \text{ or } \frac{1}{\Gamma(\alpha)}\int_0^{a/y} (1-t)^{\alpha-1} t^\xi dt,$$

$$(8.12) \qquad J^-_{\eta,\alpha}\phi_a = \frac{1}{\Gamma(\alpha+1)}\left\{\left(1-\frac{y}{a}\right)^\alpha \left(\frac{y}{a}\right)^{\eta-1} + (\eta-1)\int_{y/a}^1 (1-t)^\alpha t^{\eta-2} dt\right\} \text{ or } 0$$

for $0 < y < a$ or $a < y < \infty$ respectively; therefore, in these open intervals, $I^+_{\xi,\alpha}\phi_a$ and $J^-_{\eta,\alpha}\phi_a$ certainly exist and are continuous, also when we replace α by a purely imaginary number β, and $I^+_{\xi,\alpha}\phi_a \to I^+_{\xi,\beta}\phi_a$, $J^-_{\eta,\alpha}\phi_a \to J^-_{\eta,\beta}\phi_a$ as $\alpha \to \beta$. Hence, for any step-function f, $I^+_{\xi,\beta}f$ and $J^-_{\eta,\beta}f$ exist and are continuous almost everywhere in $(0, \infty)$, and $I^+_{\xi,0}f \equiv J^-_{\eta,0}f \equiv f$, as we can easily deduce from (8.11) and (8.12). The following theorem holds:

THEOREM V. *Let $1 \leqq p < \infty$ and $\Re(\zeta) > -1/p' = 1 - 1/p$, $\Re(\eta) > -1/p$, let*

$\Re(\lambda) \geqq 0$, and let $f(x)$ be any step-function. Then, as $\alpha \to \lambda$,

$$|I^+_{\zeta,\alpha} f - I^+_{\zeta,\lambda} f|_p \to 0, \tag{8.21}$$

$$|J^-_{\eta,\alpha} f - J^-_{\eta,\lambda} f|_p \to 0. \tag{8.22}$$

To prove (8.21), we simply take $f = \phi_a$ and $|\alpha - \lambda| < 1$. Then

$$\int_0^\infty |I^+_{\zeta,\alpha}\phi_a - I^+_{\zeta,\lambda}\phi_a|^p dy = \int_0^a \left| \frac{\Gamma(\zeta+1)}{\Gamma(\zeta+\alpha+1)} - \frac{\Gamma(\zeta+1)}{\Gamma(\zeta+\lambda+1)} \right|^p dy + \int_a^\infty |\psi(\alpha, y) - \psi(\lambda, y)|^p dy = V_1 + V_2,$$

say, where

$$\psi(\alpha, y) = \frac{1}{\Gamma(\alpha)} \int_0^{a/y} (1-t)^{\alpha-1} t^\zeta dt. \tag{8.3}$$

Obviously $V_1 \to 0$ as $\alpha \to \lambda$. When $a < y < \infty$ and y is fixed, then $\psi(\alpha, y) \to \psi(\lambda, y)$ by the Lebesgue convergence theorem, since $|(1-t)^{\alpha-1} t^\zeta| \leqq (1-a/y)^{-1} t^{\Re(\zeta)}$ and $t^{\Re(\zeta)} \subset L_1(0, 1)$. To prove $V_2 \to 0$ we need only show that $|\psi(\alpha, y)| \leqq U(y)$, where $U(y)$ does not depend on α and belongs to $L^p(a, \infty)$.

For $2a < y < \infty$, we have $1 - a/y > 1/2$,

$$|\psi(\alpha, y)| \leqq \frac{2}{|\Gamma(\alpha)|} \int_0^{a/y} t^{\Re(\zeta)} dt \leqq K y^{-\Re(\zeta)-1} = U_0(y) \subset L_p(2a, \infty).$$

For $a < y < 2a$, we have

$$\psi(\alpha, y) = \frac{1}{\Gamma(\alpha)} \int_0^{1/2} + \frac{1}{\Gamma(\alpha)} \int_{1/2}^{a/y} = \psi_1(\alpha, y) + \psi_2(\alpha, y),$$

$$|\psi_1(\alpha, y)| \leqq \frac{2}{|\Gamma(\alpha)|} \int_0^{1/2} t^{\Re(\zeta)} dt = C_1 = U_1(y) \subset L_p(a, 2a).$$

$$|\psi_2(\alpha, y)| = \frac{1}{|\Gamma(\alpha+1)|} \left| 2^{-\alpha-\zeta} - \left(1 - \frac{a}{y}\right)^\alpha \left(\frac{a}{y}\right)^\zeta + \zeta \int_{1/2}^{a/y} (1-t)^\alpha t^{\zeta-1} dt \right| < C_2 = U_2(y) \subset L_p(a, 2a). \tag{8.3.0}$$

Applying Lebesgue's theorem again, we have $V_2 \to 0$, which completes the proof.

The proof of (8.22) is similar. Let $|\alpha - \lambda| < 1$ again. We have to take into consideration that $J^-_{\eta,\alpha}\phi_a$ is bounded uniformly in α and y for $a/2 < y < a$. Furthermore, for $0 < y < a/2$ we have to replace (8.12) by

$$J^{-}_{\eta,\alpha}\phi_a = \frac{2^{1-\eta-\alpha}}{\Gamma(\alpha+1)} + \frac{\eta-1}{\Gamma(\alpha+1)}\int_{1/2}^{1}(1-t)^{\alpha}t^{\eta-2}dt$$

(8.12.0)

$$+\frac{1}{\Gamma(\alpha)}\int_{y/a}^{1}(1-t)^{\alpha-1}t^{\eta-1}dt = W_1 + W_2 + W_3,$$

where $|W_1| \leqq K$, $|W_2| \leqq K$, $|W_3| \leqq Ky^{\eta_1} \subset L_p(0, a/2)$, when $-1/p < \eta_1 < \min(0, \Re(\eta))$. For $2 \leqq p < \infty$, the theorem also follows from Theorem IIb (§9). Theorem V holds also when we suppose that ζ and η are not fixed, but that $\zeta \to \zeta_0$ and $\eta \to \eta_0$ as $\alpha \to \lambda$, where $\Re(\zeta_0) > -1/p'$ and $\Re(\eta_0) > -1/p$.

Now for $1 \leqq p \leqq \infty$ and $\Re(\zeta) > -1/p'$, $\Re(\eta) > -1/p$, we have(6)

$$|I^{+}_{\zeta,\alpha}g|_p \leqq \frac{\Gamma\{\Re(\alpha)\}\Gamma\{\Re(\zeta)+1/p'\}}{|\Gamma(\alpha)|\cdot\Gamma\{\Re(\zeta+\alpha)+1/p'\}}|g|_p;$$

(8.4)

$$|J^{-}_{\eta,\alpha}g|_p \leqq \frac{\Gamma\{\Re(\alpha)\}\Gamma\{\Re(\eta)+1/p\}}{|\Gamma(\alpha)|\cdot\Gamma\{\Re(\eta+\alpha)+1/p\}}|g|_p.$$

When in (8.21) and (8.22) $\Re(\lambda)$ is greater than zero, then $|I^{+}_{\zeta,\alpha}g|_p$ and $|J^{-}_{\eta,\alpha}g|_p$ are bounded uniformly in α for $|\alpha-\lambda| < \frac{1}{2}\Re(\lambda)$; by approximating to $g(x)$ by a sequence of step-functions we see that, for $\Re(\lambda) > 0$ and $1 \leqq p < \infty$, Theorem V is valid for any $g \subset L_p$. It is an open question whether or not it holds for $\Re(\lambda) = 0$ also, when $1 \leqq p < 2$ or $2 < p < \infty$ (cf. Theorem II), but it is certainly true in the following sense for $\lambda = 0$:

THEOREM VI. *Let $1 \leqq p < \infty$ and $f(x) \subset L_p$, let $\Re(\zeta) > -1/p'$ and $\Re(\eta) > -1/p$; let Θ be any positive number smaller than $\pi/2$, and let $\alpha \to 0$ with the restriction $|\arg\alpha| \leqq \Theta$. Then*

$$|I^{+}_{\zeta,\alpha}f - f(y)|_p \to 0; \qquad |J^{-}_{\eta,\alpha}f - f(y)|_p \to 0.$$

The proof is an immediate consequence of Theorem V, since in L_p the operations $I^{+}_{\zeta,\alpha}$ and $J^{-}_{\eta,\alpha}$ are uniformly bounded for $|\arg\alpha| \leqq \Theta$, when $|\alpha| < K$: Let $\alpha = \alpha_1 + i\alpha_2$; then $\alpha_1 > 0$, $|\alpha_2| < \alpha_1 \operatorname{tg}\Theta$, and

$$\frac{\Gamma\{\Re(\alpha)\}}{|\Gamma(\alpha)|} = \frac{\Gamma(\alpha_1+1)}{|\Gamma(\alpha+1)|}\left|1 + i\frac{\alpha_2}{\alpha_1}\right| \leqq \frac{\Gamma(\alpha_1+1)}{|\Gamma(\alpha+1)|}(1+\operatorname{tg}\Theta) < K.$$

Therefore, by (8.4), the operations have the desired property.

COROLLARY. *Let α be restricted as in Theorem* VI *and let $f \subset L_p$; then*

$$|y^{-\alpha}f^{+}_{\alpha}(y) - f(y)|_p \to 0, \quad |y^{-\alpha}f^{-}_{\alpha}(y) - f(y)|_p \to 0$$

for $1 < p < \infty$ and $1 \leqq p < \infty$ respectively, as α tends to zero.

(6) Cf. our paper cited in Footnote 4.

The case $p=1$ is not included for $f_\alpha^+(y)$ by this theorem; some results for $f_\alpha^+(y)$ $[\alpha\to 0]$ under the hypothesis $f\subset L_1(0, A)$ were given by Hardy-Littlewood (loc. cit.) and by J. D. Tamarkin(7).

9. We can state some better theorems for $p>2$. Let $2<p\leqq\infty$ and let $\mathfrak{M}_p$ be the set of all functions $f\subset L_p$ which possess a Mellin transform $F(\tau)=Mf$ in the well defined sense that $f(x)$ is representable in the form

$$f(x)=\operatorname*{index\ p\ \ l.i.m.}_{a,b\to\infty}\frac{1}{2\pi}\int_{-a}^{b}F(\tau)x^{-1/p-i\tau}d\tau=M^{-1}F, \tag{9.1}$$

where $F(\tau)\subset L_{p'}(-\infty,\ \infty)$. In consequence of the well known theory of Fourier transforms, $M^{-1}F$ is a bounded linear transformation from $L_{p'}(-\infty,\ \infty)$ into $L_p(0,\ \infty)$ and $\mathfrak{M}_p$ a subspace of L_p and smaller than L_p. When $p=2$ obviously $\mathfrak{M}_p=L_2$(8).

LEMMA 3. *Let $2<p\leqq\infty$ and $\Re(\zeta)>-1/p'$ and $\Re(\eta)>-1/p$, and let $f\subset\mathfrak{M}_p$. Then $I^+_{\zeta,\alpha}f$ and $J^-_{\eta,\alpha}f$ belong to $\mathfrak{M}_p$ also, and*

$$MI^+_{\zeta,\alpha}f=\omega(\tau,\alpha)Mf; \tag{9.21}$$

$$MJ^-_{\eta,\alpha}f=\chi(\tau,\alpha)Mf, \tag{9.22}$$

where

$$\omega(\tau,\alpha)=\frac{\Gamma(\zeta+1/p'-i\tau)}{\Gamma(\zeta+\alpha+1/p'-i\tau)};\quad \chi(\tau,\alpha)=\frac{\Gamma(\eta+1/p+i\tau)}{\Gamma(\eta+\alpha+1/p+i\tau)}. \tag{9.3}$$

We shall only outline the proof. Let $F_n(\tau)=F(\tau)$ in $(-n,\ n)$, $F_n(\tau)=0$ for $|\tau|>n$, and let $f(x,\ n)=M^{-1}F_n$. Then, for $0<y<\infty$,

$$\begin{aligned}I^+_{\zeta,\alpha}\{f(x,\ n)\}&=\frac{y^{-\zeta-\alpha}}{2\pi\Gamma(\alpha)}\int_0^y(y-x)^{\alpha-1}x^{\zeta}dx\int_{-n}^{n}F(\tau)x^{-1/p-i\tau}d\tau\\&=\frac{1}{2\pi}\int_{-n}^{n}F(\tau)y^{-1/p-i\tau}d\tau\cdot\frac{1}{\Gamma(\alpha)}\int_0^1(1-\xi)^{\alpha-1}\xi^{-1/p-i\tau}d\xi\\&=M^{-1}\{F_n(\tau)\omega(\tau,\alpha)\},\end{aligned}$$

and so (9.21) follows by $\int|F-F_n|^{p'}d\tau\to 0$, $|f(x)-f(x,\ n)|_p\to 0$ $[n\to\infty]$, and by (8.4).

---

(7) Annals of Mathematics, (2), vol. 31 (1930), pp. 219–228.

(8) When $2<p<\infty$, then (9.1) implies

$$F(\tau)=\operatorname*{index\ p'\ \ l.i.m.}_{a\to\infty}\int_{1/a}^{a}f(x)x^{-1/p'+i\tau}dx$$

in consequence of the Hille-Tamarkin theorem, vide Bulletin of the American Mathematical Society, vol. 39 (1933), pp. 768–774.

For Lemma 3 see also our paper cited above.

By (9.21) and (9.22) $I^+_{\xi,\alpha}f$ and $J^-_{\eta,\alpha}f$ are defined also when we replace α by β; for $\omega(\tau, \beta)$ and $\chi(\tau, \beta)$ are bounded for $-\infty<\tau<\infty$, therefore $\omega(\tau, \beta)Mf = \omega(\tau, \beta)F(\tau)$ and $\chi(\tau, \beta)Mf=\chi(\tau, \beta)F(\tau)$ belong to $L_{p'}(-\infty, \infty)$. Also to every $g\subset\mathfrak{M}_p$ corresponds a uniquely determined function $f\subset\mathfrak{M}_p$ or $\phi\subset\mathfrak{M}_p$ such that $I^+_{\xi,\beta}f=g$ or $J^-_{\eta,\beta}\phi=g$ respectively (cf. (6.1) and (7.2)). By the same reasoning as in §5 and §7, we have the theorems:

THEOREM IIb. *Let $2\leqq p\leqq\infty$, let $\Re(\zeta)>-1/p'$, $\Re(\eta)>-1/p$, and let $f\subset\mathfrak{M}_p$. Then*

$$|I^+_{\xi,\alpha}f - I^+_{\xi,\lambda}f|_p\to 0; \qquad |J^-_{\eta,\alpha}f - J^-_{\eta,\lambda}f|_p\to 0$$

as α tends to λ, where $\Re(\lambda)\geqq 0$.

LEMMA 4. *The operator $X_\beta f$, defined by (7.1), exists in $L_p(0, \infty)$ with both domain and range $\mathfrak{M}_p$ when $2\leqq p\leqq\infty$. So does $Y_\beta f$ when $2\leqq p<\infty$.*

THEOREM IVb. *Under the restrictions of Lemma 4, the transformations X_β or Y_β form a group in L_p, and $|X_\beta - X_{\beta_0}|_p\to 0$, $|Y_\beta - Y_{\beta_0}|_p\to 0$ as $\beta\to\beta_0$.*

10. **Characteristic values.** From (7.1) and (9.2) we easily have

$$MX_\beta f = \frac{\Gamma(1/p' - \beta - i\tau)}{\Gamma(1/p' - i\tau)}F(\tau - i\beta) \qquad [2\leqq p\leqq\infty], \tag{10.11}$$

$$MY_\beta f = \frac{\Gamma(1/p + i\tau)}{\Gamma(1/p + \beta + i\tau)}F(\tau - i\beta) \qquad [2\leqq p<\infty], \tag{10.12}$$

where $f=M^{-1}F\subset\mathfrak{M}_p$. We shall now deal with the equations

$$X_\beta f = lf, \tag{10.21}$$

$$Y_\beta f = lf, \tag{10.22}$$

$$X_\beta f = l\bar{f}. \tag{10.23}$$

Let $\beta=i\rho$, where ρ is real. Obviously (10.21) is equivalent to

$$h(\tau + \rho) = lh(\tau);\ F(\tau) = h(\tau)/\Gamma(1/p' - i\tau)\subset L_{p'}(-\infty, \infty), \tag{10.31}$$

and (10.22) or (10.23) to

$$k(\tau + \rho) = lk(\tau);\ F(\tau) = k(\tau)\Gamma(1/p + i\tau)\subset L_{p'}(-\infty, \infty) \tag{10.32}$$

or

$$h(\tau + \rho) = l\bar{h}(-\tau);\ F(\tau) = h(\tau)/\Gamma(1/p' - i\tau)\subset L_{p'}(-\infty, \infty). \tag{10.33}$$

By means of the well known asymptotic expansion of $|\Gamma(\sigma+i\tau)|$ for $t\to\pm\infty$ we arrive at the following results:

The transformation $X_\beta f$ $[\beta\neq 0]$ with domain $\mathfrak{M}_p$ $[2\leqq p\leqq\infty]$ has no characteristic values at all.

The characteristic values of $Y_\beta f$ $[\beta \neq 0$, *domain* $\mathfrak{M}_p$, $2 \leqq p < \infty]$ *are the set of points* l *for which* $\exp(-\frac{1}{2}\pi|\beta|) < |l| < \exp(\frac{1}{2}\pi|\beta|)$. *To every characteristic value* l *corresponds an infinity of characteristic functions.*

We construct all these functions by (10.32), taking $k(\tau)$ in $(0, |\rho|)$ as an arbitrary function belonging to $L_{p'}(0, |\rho|)$.

Also, from the group property of Y_β, we can deduce the result:

Let $\Re(\kappa) = \Re(\lambda) = 0$, *let* κ/λ *be no rational number, and let* $Y_\kappa f = f$ *and* $Y_\lambda f = f$ *and* $f \subset \mathfrak{M}_p$. *Then* $f(x) = ce^{-x}$, *and* $Y_\gamma f = f$ *for any* γ *such that* $\Re(\gamma) \geqq 0$.

Furthermore, by (10.33), we can prove the result:

The equation (10.23) *has a solution* $f \subset \mathfrak{M}_p$ *if, and only if,* $|l| = 1$. *To every number* l *of this kind corresponds an infinity of solutions* $f \subset \mathfrak{M}_p$.

For instance, $f(x) = e^{-1/x}x^{-\beta-1}$ is a solution of (10.23) for $l = 1$.

11. It is not difficult to show that $I^+_{\zeta,\beta}f$ and $J^-_{\eta,\alpha}f$ certainly exist and are integrable for any $f \subset L_p$ when $\Re(\zeta) > -1/2$ and $\Re(\eta) > -1/p$ and $2 < p < \infty$. We have

THEOREM VII. *Let* $2 < p < \infty$ *and* $f(x) \subset L_p$; *let* $\Re(\zeta) > -1/2$, $\Re(\eta) > -1/p$. *Then there exist functions* $I^+_{\zeta,\beta}f$ *and* $J^-_{\eta,\alpha}f$, *defined in* $(0, \infty)$ *and such that, for any positive finite number* A, $I^+_{\zeta,\beta}f$ *and* $J^-_{\eta,\beta}f$ *belong to* $L_2(0, A)$ *and that*

$$\int_0^A |I^+_{\zeta,\alpha}f - I^+_{\zeta,\beta}f|^2 dy \to 0; \quad \int_0^A |J^-_{\eta,\alpha}f - J^-_{\eta,\beta}f|^2 dy \to 0 \quad [\alpha \to \beta].$$

Proof. Let $f_n(x) = f(x)$ or $f_n(x) = 0$ for $0 \leqq x \leqq n$ or $x > n$ respectively $[n = 1, 2, \cdots]$. Then $f_n(x) \subset L_2(0, \infty)$, and by Theorem II, the function $g_\alpha(y, n) = I^+_{\zeta,\alpha}f_n$ belongs to $L_2(0, \infty)$ and converges strongly to a function $g_\beta(y, n)$ as $\alpha \to \beta$,

$$|g_\alpha(y, n) - g_\beta(y, n)|_2 \to 0. \tag{11.1}$$

We now define $I^+_{\zeta,\beta}f$ in $(0, \infty)$ by putting

$$I^+_{\zeta,\beta}f = g_\beta(y, n) \quad \text{for} \quad n - 1 < y \leqq n \qquad [n = 1, 2, \cdots]. \tag{11.2}$$

This function then has the desired properties. For $0 < y \leqq n$ obviously

$$g_\alpha(y, n) = I^+_{\zeta,\alpha}f,$$

and so

$$\int_0^n |I^+_{\zeta,\alpha}f - g_\beta(y, n)|^2 dy \leqq \int_0^\infty |g_\alpha(y, n) - g_\beta(y, n)|^2 dy \to 0 \tag{11.3}$$

as α tends to β. Therefore, for $1 \leqq m < n$,

$$\left(\int_0^m |g_\beta(y,n)-g_\beta(y,m)|^2dy\right)^{1/2} \leqq \left(\int_0^m |I^+_{\xi,\alpha}f-g_\beta(y,m)|^2dy\right)^{1/2}$$
$$+\left(\int_0^n |g_\beta(y,n)-I^+_{\xi,\alpha}f|^2dy\right)^{1/2}\to 0 \qquad [\alpha\to\beta].$$

Hence $g_\beta(y,m)\equiv g_\beta(y,n)$ in $(0,m)$, and

$$I^+_{\xi,\beta}f \equiv g_\beta(y,n) \tag{11.4}$$

in $(0,n)$ for $n=1,2,\cdots$.

Therefore $I^+_{\xi,\beta}f\subset L_2(0,n)$ for any n, and, by (11.3) and (11.4),

$$\int_0^n |I^+_{\xi,\alpha}f-I^+_{\xi,\beta}f|^2dy=\int_0^n |I^+_{\xi,\alpha}f-g_\beta(y,n)|^2dy\to 0$$

as $\alpha\to\beta$; which proves the first part of the theorem.

Let us take (cf. 4.4) $0<y\leqq n$,

$$J^-_{\alpha,\eta}f=\frac{y^\eta}{\Gamma(\alpha)}\left\{\int_y^n+\int_n^\infty\right\}(x-y)^{\alpha-1}x^{-\eta-\alpha}f(x)dx \tag{11.5}$$
$$=\phi_\alpha(y,n)+\psi_\alpha(y,n).$$

Plainly $\phi_\alpha(y,n)=J^-_{\alpha,\eta}f_n$ in $(0,n)$. Now $f_n\subset L_2(0,\infty)$ and $\Re(\eta)>-1/p>-1/2$, and so, by Theorem II, $J^-_{\beta,\eta}f_n$ exists, belongs to L_2, and

$$\int_0^n |\phi_\alpha(y,n)-J^-_{\eta,\beta}f_n|^2dy\leqq\int_0^\infty |J^-_{\eta,\alpha}f_n-J^-_{\eta,\beta}f_n|^2dy\to 0 \quad [\alpha\to\beta]. \tag{11.6}$$

Let b be any positive number smaller than n. Then there exists a function $\psi_\beta(y,n)$, depending on n but not on b, such that

$$\int_0^b |\psi_\alpha(y,n)-\psi_\beta(y,n)|^2dy\to 0 \qquad [\alpha\to\beta], \tag{11.7}$$

as we shall now show. For $0<y<n$ and $n\leqq x<\infty$, the function

$$\chi(x,y,\alpha)=\{\Gamma(\alpha)\}^{-1}y^{-\eta}(x-y)^{\alpha-1}x^{-\eta-\alpha}f(x)$$

tends to $\chi(x,y,\beta)$ as $\alpha\to\beta$, and

$$|\chi(x,y,\alpha)|<Kx^{-\Re(\eta)-1}|f(x)|\subset L_1(n,\infty)$$

when y is fixed and $|\alpha|<K$. Hence, by Lebesgue's convergence theorem, $\psi_\beta(y,n)$ exists, and $\psi_\alpha(y,n)-\psi_\beta(y,n)\to 0$ when $0<y<n$. For $0<y\leqq b<n$, we have $x-y\geqq(1-b/n)x=cx$,

$$|\psi_\alpha(y,n)|\leqq|\Gamma(\alpha)|^{-1}y^{\Re(\eta)}\int_n^\infty c^{\Re(\alpha)-1}x^{-\Re(\eta)-1}|f(x)|\,dx\leqq Ky^{\Re(\eta)},$$

where K does not depend on α when $|\alpha-\beta|<1$. Hence also

$$|\psi_\alpha(y, n) - \psi_\beta(y, n)| \leqq K y^{\Re(\eta)}.$$

Since $y^{\Re(\eta)} \subset L_2(0, b)$, applying Lebesgue's theorem again, we obtain (11.7). In consequence of (11.5)–(11.7), for $h_\beta(y, n) = \phi_\beta(y, n) + \psi_\beta(y, n)$ we have

$$\int_0^b |J^-_{\eta,\alpha} f - h_\beta(y, n)|^2 dy \to 0 \qquad [\alpha \to \beta]. \tag{11.8}$$

When we define $J^-_{\eta,\beta} f$ in $(0, \infty)$ by

$$J^-_{\eta,\beta} f = h_\beta(y, n) \qquad \text{for } n - 1 \leqq y < n \ [n = 1, 2, \cdots],$$

then, by the argument applied in (11.4), it easily follows that $J^-_{\eta,\beta} f \equiv h_\beta(y, n)$ for $0 < y < n$, which completes the proof.

THE UNIVERSITY,
EDGBASTON, BIRMINGHAM, ENGLAND.

THE STRUCTURE OF THE GROUP RING OF A p-GROUP OVER A MODULAR FIELD

BY

S. A. JENNINGS

This paper deals with the group ring of a group of prime power order over the prime field $GF(p)$, where p is the prime dividing the order of the group. It is well known that in the case of the group ring of a group over a field whose characteristic divides the order of the group, the ordinary theory of group characters is no longer valid: recently, Brauer and Nesbitt(1) have investigated the properties of the modular representations in this case, but this general theory yields only little in the special case that we consider here. We investigate the group ring from the point of view of the structure of its radical, and in particular, determine a basis for, and the ranks of, the various powers of the radical in terms of the elements and order of a new series of characteristic subgroups. These subgroups are defined by a certain minimal property which combines the commutator and the pth power structure of the group, and should prove useful in general investigations on the structure of p-groups.

1. It is well known that the group ring of a group of order g is semi-simple, provided the characteristic of the underlying field is zero, or a prime which does not divide g(2). If, however, the underlying field has characteristic p, and p divides g, then it is readily seen that the group ring has a radical which is not zero. Let the elements of the group be $G_1=1, G_2, \cdots, G_g$. Consider the element $\sigma=G_1+G_2+\cdots+G_g$ in the group ring. We have $\sigma\cdot G_i=\sigma$, and hence, if $A=\sum\alpha_iG_i$ is any element in the ring, $\sigma\cdot A=A\cdot\sigma=(\sum\alpha_i)\cdot\sigma$; that is, scalar multiples of σ form an ideal (σ). However, $(\sigma)\cdot(\sigma)=0$, since $\sigma\cdot\sigma=\sigma\cdot(\sum 1)=0$, $((\sum 1)=g\equiv 0$ modulo $p)$, and hence the group ring contains a nilpotent ideal different from zero. We have proved(3), therefore,

THEOREM 1.1. *The group ring of a group over a field whose characteristic divides the order of the group is not semi-simple.*

We investigate the structure of the group ring in the extreme case, where

Presented to the Society, October 28, 1939; received by the editors November 24, 1939, and, in revised form, September 20, 1940.

The essentials of this paper comprise a thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy at the University of Toronto, May, 1939. It was written under the direction of Professor Richard Brauer, and I would like to acknowledge my debt to him for the assistance and encouragement that he gave me.

(1) R. Brauer and C. Nesbitt, University of Toronto Studies, Mathematical Series, no. 4.

(2) See, for example, van der Waerden, *Moderne Algebra*, vol. 2, §125.

(3) E. Noether, Mathematische Zeitschrift.

the order of the group is a power of the characteristic of the underlying field. In what follows, $\mathfrak{G}$ is a p-group of order p^s, and Γ is the group ring of $\mathfrak{G}$ over the prime field $GF(p)$. Where necessary, we emphasize the fact that Γ is the group ring of $\mathfrak{G}$ by writing $\Gamma=\Gamma(\mathfrak{G})$.

We prove first the following

THEOREM 1.2. *The radical, $\mathfrak{N}$, of the group ring Γ of a p-group $\mathfrak{G}$ of order p^s over the field $GF(p)$ is of rank p^s-1, and has as basis all elements of the form G_i-1, where $G_i\in\mathfrak{G}$, $G_i\neq 1$. A necessary and sufficient condition that an element $A=\sum\alpha_iG_i$ of Γ lie in $\mathfrak{N}$ is that $\sum\alpha_i=0$. The semi-simple part of Γ is of rank one and is isomorphic to $GF(p)$.*

It is known that the only irreducible modular representation of $\mathfrak{G}$ is the 1-representation(4) and hence, if $\mathfrak{A}$ is any representation of $\mathfrak{G}$, $\mathfrak{A}$ may be transformed to the form

$$\mathfrak{A}\sim\begin{pmatrix}(1)&&\\&\ddots&\\&&(1)\end{pmatrix}.$$

Every representation of $\mathfrak{G}$ also gives a representation of Γ. In particular, if we take for $\mathfrak{A}$ the regular representation of $\mathfrak{G}$, we obtain a (1:1) representation of Γ. Every element of the form (G_i-1), where $G_i\in\mathfrak{G}$, is represented by matrices

$$G_i-1\leftrightarrow\begin{pmatrix}0&&\\&\ddots&\\&&0\end{pmatrix} \tag{1.3}$$

with zeros in the main diagonal. All elements which are represented by matrices of this form with zeros in and above the main diagonal form a nilpotent ideal, and hence G_i-1 belongs to the radical $\mathfrak{N}$ of Γ. Since there are p^s-1 independent elements G_i-1, the rank of $\mathfrak{N}$ is at least p^s-1. Since Γ is of rank p^s, $\Gamma/\mathfrak{N}$ is of rank 0 or 1. The element 1 is not in $\mathfrak{N}$, however, since it is not nilpotent; and hence the rank of Γ is p^s-1, and the rank of $\Gamma/\mathfrak{N}$ is 1. It follows that $\Gamma/\mathfrak{N}\simeq GF(p)$. That the elements G_i-1 form a basis for $\mathfrak{N}$ follows, since there are p^s-1 of them, and they are independent. The remainder of the theorem is immediate(5).

COROLLARY 1.4. *If G is any element of $\mathfrak{G}$, then $G\equiv 1$ modulo $\mathfrak{N}$.*

The corollary follows from the fact that $G-1\equiv 0$ modulo $\mathfrak{N}$.

(4) L. E. Dickson, these Transactions, vol. 8 (1907), pp. 389–398. Cf. also Brauer and Nesbitt, loc. cit.

(5) L. Lombardo-Radici, in a recent paper has given a group-theoretical proof of Theorem 1.2.

2. When $\mathfrak{N}$ is the radical of Γ, we may form the various powers of $\mathfrak{N}$([6])

$$\mathfrak{N} \supset \mathfrak{N}^2 \supset \cdots \supset \mathfrak{N}^L \supset \mathfrak{N}^{L+1} = 0. \tag{2.1}$$

$\mathfrak{N}^w$ consists of sums of products of w elements of $\mathfrak{N}$. If the situation is as in (2.1) with $\mathfrak{N}^L \neq 0$, $\mathfrak{N}^{L+1}=0$, we say that L is the *exponent* of $\mathfrak{N}$. We define $\mathfrak{N}^0=\Gamma$.

Let $\mathfrak{K}_\lambda$ be the set of all elements $K_\lambda \in \mathfrak{G}$ such that $K_\lambda \equiv 1$ modulo $\mathfrak{N}^\lambda$. $\mathfrak{K}_\lambda$ is a subgroup of $\mathfrak{G}$, since if K_λ, $K_\lambda' \equiv 1$ modulo $\mathfrak{N}^\lambda$, then $K_\lambda K_\lambda' \equiv 1$ modulo $\mathfrak{N}^\lambda$. Moreover, $\mathfrak{K}_\lambda$ is a self-conjugate subgroup of $\mathfrak{G}$, for if G is any element of $\mathfrak{G}$ then $G^{-1}K_\lambda G \equiv 1$ modulo $\mathfrak{N}^\lambda$, since $\mathfrak{N}^\lambda$ is an ideal of Γ. Indeed, it is clear that $\mathfrak{K}_\lambda$ is a characteristic subgroup of $\mathfrak{G}$, since any automorphism of $\mathfrak{G}$ leaves $\mathfrak{N}$ and its powers unaltered. Since $\mathfrak{N}^\lambda \supset \mathfrak{N}^\mu$ for $\lambda < \mu$, we have $\mathfrak{K}_\lambda \supseteq \mathfrak{K}_\mu$, $\lambda < \mu$. By Corollary 1.4 $\mathfrak{K}_1 = \mathfrak{G}$, and by (2.1) $\mathfrak{K}_{L+1}=1$. We have proved, therefore,

THEOREM 2.2. *The sets $\mathfrak{K}_\lambda$, $\lambda = 1, 2, \cdots$, consisting of group elements which may be written in the form $1+n_\lambda$, $n_\lambda \in \mathfrak{N}^\lambda$, form a decreasing series of characteristic subgroups of $\mathfrak{G}$:*

$$\mathfrak{G} = \mathfrak{K}_1 \supseteq \mathfrak{K}_2 \supseteq \cdots \supseteq \mathfrak{K}_{L+1} = 1.$$

We shall refer to these subgroups as the $\mathfrak{K}$-series of the group $\mathfrak{G}$([7]). We write $\mathfrak{K}_\lambda(\mathfrak{G})$ for $\mathfrak{K}_\lambda$ when necessary to stress the fact that $\mathfrak{K}_\lambda$ is a member of the $\mathfrak{K}$-series of a particular group $\mathfrak{G}$.

THEOREM 2.3. *The $\mathfrak{K}$-series of any group $\mathfrak{G}$ has the following properties*

(1) $(\mathfrak{K}_\lambda, \mathfrak{K}_\mu) \subseteq \mathfrak{K}_{\lambda+\mu}$([8]),

(2) $K_i^p \in \mathfrak{K}_{ip}$ *if* $K_i \in \mathfrak{K}_i$.

As a consequence of (1) *and* (2), *we have*

(3) $\mathfrak{K}_\lambda/\mathfrak{K}_{\lambda_1}$ *is abelian of type* $(p, p, \cdots, p)$.

Proof. To prove (1) we must show that if $K_\lambda \in \mathfrak{K}_\mu$, $K_\mu' \in \mathfrak{K}_\mu$, then $K_\lambda^{-1}K_\mu'^{-1}K_\lambda K_\mu' \in \mathfrak{K}_{\lambda+\mu}$. Now if $K_\lambda \in \mathfrak{K}_\lambda$, $K_\mu' \in \mathfrak{K}_\mu$, we may write

$$\begin{aligned} K_\lambda &= 1 + n_\lambda, & n_\lambda &\in \mathfrak{N}^\lambda, \\ K_\mu' &= 1 + n_\mu', & n_\mu' &\in \mathfrak{N}^\mu, \end{aligned} \tag{2.4}$$

and we have

([6]) Cf. Dickson, *Algebren und ihre Zahlentheorie*, chap. 6.

([7]) In a paper which appeared after the present paper had been written, H. Zassenhaus has described (*Ein Verfahren, jeder endlichen p-Gruppe einen Lie-Ring mit Charakteristik p zuzuordnen*, Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität, vol. 13 (1939), pp. 200–206), for a given p-group, a series which he calls "the dimensional groups modulo p." This series appears to be identical with our $\mathfrak{K}$-series. It is interesting that these groups should arise both from the ring, and from the Lie-algebra associated with a p-group over a modular field.

([8]) For this notation, see P. Hall, *A contribution to the theory of groups of prime power orders*, Proceedings of the London Mathematical Society, vol. 36 (1933–1934), pp. 29–95, especially §2.

$$K_\lambda^{-1}K_\mu'^{-1}K_\lambda K_\mu = 1 + K_\lambda^{-1}K_\mu'^{-1}(K_\lambda K_\mu' - K_\mu' K_\lambda)$$
$$= 1 + K_\lambda^{-1}K_\mu'^{-1}((1+n_\lambda)(1+n_\mu') - (1+n_\mu')(1+n_\lambda))$$
$$= 1 + K_\lambda^{-1}K_\mu'^{-1}(n_\lambda n_\mu' - n_\mu' n_\lambda) = 1 + \bar{n}_{\lambda+\mu}$$

where $\bar{n}_{\lambda+\mu}\in\mathfrak{N}^{\lambda+\mu}$, which proves that $(K_\lambda, K_\mu')\equiv 1$ modulo $\mathfrak{N}^{\lambda+\mu}$. Again, if $K_i\in\mathfrak{K}_i$, then $K_i=1+n_i$, where $n_i\in\mathfrak{N}^i$; and since 1 and n_i permute, and we are in a field of characteristic p, we have

$$K_i^p = (1+n_i)^p = 1 + n_i^p,$$

which shows that $K_i^p\equiv 1$ modulo $\mathfrak{N}^{ip}$, and proves (2). Statement (3) follows readily from (1) and (2).

Theorem 2.4, (1), shows that the $\mathfrak{K}$-series is a central series of $\mathfrak{G}$, as defined by Hall(9).

By 2.4, (3), $\mathfrak{K}_\lambda/\mathfrak{K}_{\lambda+1}$ is elementary abelian. Let $\mathfrak{K}_\lambda/\mathfrak{K}_{\lambda+1}$ be of order p^{d_λ} (it may happen that $d_\lambda=0$ if $\mathfrak{K}_\lambda=\mathfrak{K}_{\lambda+1}$), and let $F_{\lambda,1}, \cdots, F_{\lambda,d_\lambda}$ be a complete set of representatives in $\mathfrak{G}$ of a minimal basis for $\mathfrak{K}_\lambda/\mathfrak{K}_{\lambda+1}$ (again if $\mathfrak{K}_\lambda=\mathfrak{K}_{\lambda+1}$ we set $F_{\lambda,i}=1$). Then any element $G\in\mathfrak{G}$ may be written

$$G = F_{1,1}^{\varepsilon_{1,1}}F_{1,2}^{\varepsilon_{1,2}}\cdots F_{1,d}^{\varepsilon_{1,d_1}}\cdots F_{\lambda,1}^{\varepsilon_{\lambda,1}}\cdots F_{\lambda,d_\lambda}^{\varepsilon_{\lambda,d_\lambda}}\cdots \tag{2.5}$$

where $0\leqq x_{\lambda,i}<p$, and the $x_{\lambda,i}$ are uniquely determined modulo p.

If $F_{\lambda,i}\neq 1$, then $F_{\lambda,i}$ is not in $\mathfrak{K}_{\lambda+1}$, and hence $(F_{\lambda,i}-1)$ is in $\mathfrak{K}_\lambda$, and not in $\mathfrak{K}_{\lambda+1}$. Suppose K_λ is any element of $\mathfrak{K}_\lambda$. Using (2.5) we may write

$$K_\lambda \equiv F_{\lambda,1}^{x_{\lambda,1}}\cdots F_{\lambda,d_\lambda}^{x_{\lambda,d_\lambda}} \text{ modulo } \mathfrak{K}_{\lambda+1}, \qquad 0\leqq x_{\lambda,i}<p,$$

where the $x_{\lambda,i}$ are uniquely determined. Using the identity

$$(AB-1) = (A-1)(B-1) + (A-1) + (B-1) \tag{2.6}$$

we readily obtain

$$K_\lambda - 1 \equiv \sum_i x_{\lambda,i}(F_{\lambda,i}-1) \text{ modulo } \mathfrak{N}^{\lambda+1} \tag{2.7}$$

(since $x_{\lambda,i}<p$, we may suppose $x_{\lambda,i}$ in the underlying field). Conversely, if $K_\lambda\in\mathfrak{K}_\lambda$, and we have a relation of the form

$$K_\lambda - 1 \equiv \sum_i y_{\lambda,i}(F_{\lambda,i}-1) \text{ modulo } \mathfrak{N}^{\lambda+1}, \qquad y_{\lambda,i}\in GF(p),$$

this implies that

$$K_\lambda \equiv F_{\lambda,1}^{y_{\lambda,1}}\cdots F_{\lambda,d_\lambda}^{y_{\lambda,d_\lambda}} \text{ modulo } \mathfrak{K}_{\lambda+1}, \qquad 0\leqq y_{\lambda,i}<p.$$

From these facts we obtain

(9) Hall, loc. cit., §2.4.

THEOREM 2.8. *A minimal generating set* $\{F_{\lambda,1}, \cdots, F_{\lambda,d_\lambda}\}$ *for* $\mathfrak{R}_\lambda$ *modulo* $\mathfrak{R}_{\lambda+1}$ *may be taken as any maximal set of elements* $F_{\lambda,i}$ *of* $\mathfrak{R}_\lambda$ *for which* $(F_{\lambda,i}-1)$ *are linearly independent modulo* $\mathfrak{N}^{\lambda+1}$.

3. We are now in a position to determine a basis for $\mathfrak{N}^\lambda$ modulo $\mathfrak{N}^{\lambda+1}$ in terms of the elements $F_{\lambda,i}$ defined above. For fixed w consider all products of the form

$$\prod_{i,\lambda} (F_{\lambda,i}-1)^{\alpha_{\lambda,i}}, \qquad 0 \leqq \alpha_{\lambda,i} < p, \tag{3.1}$$

with $\sum_{i,\lambda}(\lambda\alpha_{\lambda,i})=w$, the summation extending over the same λ and i as in the product. We insist that in such a product those factors which are present shall be in the natural order of increasing i and λ, as in (2.5). Let us call the various distinct products (3.1)

$$N_1^{(w)}, N_2^{(w)}, \cdots, N_{l_w}^{(w)}.$$

We define w to be the "weight" of these products. Letting $w=1, 2, 3, \cdots$, we obtain exactly p^a-1 such products which are formally distinct, since $d_1+d_2+\cdots=a$. However, any element of the form $G-1$, where $G\in\mathfrak{G}$, may be expressed as a linear combination of the $N_\lambda^{(w)}$, since by (2.5) $G-1=(\prod F_{\lambda,i}^{\alpha_{\lambda,i}})-1$ and by using the identity (2.6) a sufficient number of times we get the result. Hence, using Theorem 1.2, we see that the products $N_\lambda^{(w)}$ are independent, and form a basis for $\mathfrak{N}$.

We prove now the following:

THEOREM 3.2. *The elements* $N_\lambda^{(i)}$, *with* $i\geqq w$, *form a basis for* $\mathfrak{N}^w$.

Proof. Certainly the $N_\lambda^{(i)}$ lie in $\mathfrak{N}^w$, and are independent. We have to show that every $n_w\in\mathfrak{N}^w$ can be expressed as a linear combination of these elements. Suppose the theorem false. Let w be the largest power of $\mathfrak{N}$ for which the theorem does not hold (such a largest power exists since the theorem is true for $w=L+1$), and let $n_w\in\mathfrak{N}^w$ be an element which cannot be expressed linearly in terms of the $N_\lambda^{(i)}$. Then there is no relation of the form

$$n_w \equiv \sum_\lambda c_\lambda N_\lambda^{(w)} \text{ modulo } \mathfrak{N}^{w+1}$$

where the c_λ are in the underlying field, since otherwise $n_w-\sum c_\lambda N_\lambda^{(w)}$, which is in $\mathfrak{N}^{w+1}$, could be expressed by $N_\mu^{(j)}$ with $j\geqq w+1$, by our choice of w, and we would obtain a contradiction.

Now by the definition of $\mathfrak{N}^w$, n_w may be written, modulo $\mathfrak{N}^{w+1}$, in the form

$$n_w \equiv \sum \prod_i (F_{1,\rho_i}-1) \qquad (w \text{ factors}). \tag{3.3}$$

There must be at least one product $\prod_i(F_{1,\rho_i}-1)$ which, modulo $\mathfrak{N}^{w+1}$, is linearly independent of the $N_\lambda^{(w+1)}$.

More generally, consider all products

$$\pi = \prod_i (F_{\sigma_i,\tau_i} - 1) \qquad \text{with } \sum_i \sigma_i = w. \tag{3.4}$$

A product π is here taken as "higher" than a product π' if the number of factors with $\sigma_i = w$ is greater in π than in π'; if they have the same number of such factors, then π is "higher" than π' if π has the greater number of factors with $\sigma_i = w-1$, etc. We select a highest product π (cf. (3.4)) which is linearly independent of the $N_\lambda^{(w)}$ modulo $\mathfrak{N}^{w+1}$.

We show now that if π is chosen thus, and if we interchange two consecutive factors in π, then the new product thus obtained (which is as "high" as π), again is linearly independent of the $N_\lambda^{(w)}$.

Let $(A-1)$ and $(B-1)$ be the two factors in π to be interchanged, A being in $\mathfrak{R}_\rho$ and B in $\mathfrak{R}_\sigma$. Set $\pi = \pi_1(B-1)\cdot(A-1)\pi_2$. Using the identity

$$(B-1)\cdot(A-1) = (A-1)\cdot(B-1) + (AB-1)\cdot(C-1) + (C-1) \tag{3.5}$$

where $C = (B, A)$, we obtain

$$\pi = \pi_1(A-1)\cdot(B-1) + \pi_1(AB-1)\cdot(C-1)\pi_2 + \pi_1(C-1)\pi_2.$$

Using (2.4), (1) we see that the second term on the right is in $\mathfrak{N}^{w+1}$, and the factor $(C-1)$ is in $\mathfrak{N}^{\rho+\sigma}$, whence, using (2.5) and (2.7), we get a formula of the form

$$\pi \equiv \pi_1(A-1)(B-1)\pi_2 + \sum_\mu a_\mu \pi_1(F_{\rho+\sigma,\mu} - 1)\pi_2 \text{ modulo } \mathfrak{N}^{w+1}, \qquad a_\mu \in GF(p).$$

All terms of the form $\pi_1(F_{\rho+\sigma,\mu}-1)\pi_2$ are "higher" than π, however, and can therefore be expressed by the $N_\lambda^{(w)}$, modulo $\mathfrak{N}^{w+1}$, and since π is independent of the $N_\lambda^{(w)}$, so is $\pi_1(A-1)(B-1)\pi_2$.

We may therefore take the factors in π in any order we please, and still have a product which is independent of the $N_\lambda^{(w)}$. Order the factors in π as in (3.1):

$$\pi = \prod_i (F_{i,\nu} - 1)^{\beta_{i,\nu}}, \qquad \sum_i (i\beta_{i,\nu}) = w.$$

Suppose an exponent $\beta_{i,\nu}$ were greater than $p-1$. We replace $(F_{i,\nu}-1)^p$ by $(F_{i,\nu}^p - 1)$, and since $F_{i,\nu}^p \in \mathfrak{R}_{ip}$, we use (2.7) to express $(F_{i,\nu}-1)^p$ in the form

$$(F_{i,\nu} - 1)^p \equiv \sum_\mu c_\mu (F_{ip,\mu} - 1) \text{ modulo } \mathfrak{N}^{ip+1}.$$

On making this substitution for $(F_{i,\nu}-1)^p$ in π we get π expressed as a sum of terms all "higher" than π, which is impossible, since π is independent of the $N_\lambda^{(w)}$ and these "higher" terms are not. However, if all the $\beta_{i,\nu}$ are less than p, then π itself is an $N_\lambda^{(w)}$. In any case we get a contradiction, and Theorem 3.2 is therefore proven.

As immediate corollaries we have:

THEOREM 3.6. *The elements $N_\lambda^{(w)}$ for fixed w form a basis of $\mathfrak{N}^w$ modulo $\mathfrak{N}^{w+1}$. The number l_w of these elements is equal to the rank of $\mathfrak{N}^w/\mathfrak{N}^{w+1}$.*

THEOREM 3.7. *The rank l_w of $\mathfrak{N}^w/\mathfrak{N}^{w+1}$ is equal to the coefficient of x^w in the expansion of*

$$(1 + x + x^2 + \cdots + x^{p-1})^{d_1} \cdot (1 + x^2 + x^4 + \cdots + x^{2(p-1)})^{d_2} \cdots \cdot (1 + x^\lambda + x^{2\lambda} + \cdots + x^{\lambda(p-1)})^{d_\lambda} \cdots$$

and the exponent L of $\mathfrak{N}$ is equal to $\sum_\lambda \lambda d_\lambda (p-1)$.

To prove Theorem 3.7 we need only notice that the coefficient of x^w is the number of ways of selecting the formally distinct products (3.1).

The set of numbers $(L; l_1, \cdots, l_L)$ has been called the "genus" of the radical by Hazlett[10]. They determine, to a certain extent, the structure of $\mathfrak{N}$. We have obtained an explicit expression for the genus of the radical of the group ring of a p-group in terms of the orders of the $\mathfrak{K}$-series of the group.

4. In this section we establish certain properties of the $\mathfrak{K}$-series which are necessary to identify this series abstractly. We proceed at once to prove:

THEOREM 4.1. *If $\mathfrak{H}$ is a self-conjugate subgroup of $\mathfrak{G}$, and $\mathfrak{H} \subseteq \mathfrak{K}_i(\mathfrak{G})$, then the $\mathfrak{K}$-series of $\mathfrak{G}/\mathfrak{H}$ starts with $\mathfrak{G}/\mathfrak{H}$, $\mathfrak{K}_2/\mathfrak{H}$, $\cdots$, $\mathfrak{K}_i/\mathfrak{H}$, that is,*

$$\mathfrak{K}_\lambda(\mathfrak{G}/\mathfrak{H}) \simeq \mathfrak{K}_\lambda(\mathfrak{G})/\mathfrak{H}, \qquad \lambda \leqq i.$$

Let $\mathfrak{G}/\mathfrak{H} = \mathfrak{G}'$. Then we have a homomorphic mapping $\mathfrak{G} \to \mathfrak{G}'$. Let Γ' be the group ring of $\mathfrak{G}'$ over the field $GF(p)$. The homomorphism above can be extended to a homomorphism of Γ upon Γ' in a natural manner as follows. Let $\{\overline{G}_\mu\}$ be a complete set of representatives in $\mathfrak{G}$ of $\mathfrak{G}/\mathfrak{H}$, and let H_ν, $\nu = 1, 2, \cdots$, be the elements of $\mathfrak{H}$. Every element $\gamma \in \Gamma$ may be written therefore

$$\gamma = \sum_{\mu,\nu} c_{\mu,\nu} \overline{G}_\mu H_\nu \qquad C_{\mu,\nu} \in GF(p).$$

In the mapping $\mathfrak{G} \to \mathfrak{G}'$, every element $\overline{G}_\mu H_\nu \to \overline{G}_\mu \mathfrak{H}$. A mapping of Γ upon Γ' can be defined thus:

$$\gamma = \sum c_{\mu,\nu} \overline{G}_\mu H_\nu \to \gamma' = \sum_\mu \Big(\sum_\nu c_{\mu\nu} \Big) \overline{G}_\mu \mathfrak{H}. \tag{4.2}$$

It is readily verified that the mapping as given by (4.2) is a homomorphism. Consider those elements of Γ which map into 0 in Γ': if γ is such an element, by (4.2) we must have $\sum_\nu c_{\mu,\nu} = 0$ and hence we may write

$$\gamma = \sum c_{\mu,\nu} \overline{G}_\mu (H_\nu - 1);$$

[10] O. Hazlett, *On the classification and invariantive characterization of nilpotent algebras*, American Journal of Mathematics, vol. 6 (1905), pp. 109–138. Cf. also G. Pickert, Mathematische Annalen, vol. 116 (1938), pp. 217–280.

and since $\mathfrak{H} \subseteq \mathfrak{K}_i(\mathfrak{G})$, we see that $\gamma \in \mathfrak{N}^i(\Gamma)$. Hence if γ and γ_1, in Γ, map into the same element in Γ', then

$$\gamma \equiv \gamma_1 \text{ modulo } \mathfrak{N}^i(\Gamma).$$

It is clear that $\mathfrak{N}(\Gamma)$ maps upon a subset of $\mathfrak{N}(\Gamma')$, and hence $\mathfrak{N}(\Gamma)^\rho$ maps upon a subset of $\mathfrak{N}(\Gamma')^\rho$. Conversely, if $\nu' = \sum c_\tau(\overline{G}_\tau \mathfrak{H} - 1)$, C_τ in $GF(p)$, is an element of Γ', the element $\nu = \sum c_\tau(\overline{G}_\tau - 1)$ of Γ is mapped upon ν'. It follows that if ν_ρ' is an element of $\mathfrak{N}(\Gamma')^\rho$, we may find an element ν_ρ of $\mathfrak{N}(\Gamma)^\rho$ with this image ν_ρ'. If $\hat{\nu}_\rho$ is any other element of Γ with this image, then $\nu_\rho \equiv \hat{\nu}_\rho$ modulo $\mathfrak{N}^i$, according to the remark above. If $\rho \leqq i$, then $\tilde{\nu}_\rho \equiv \nu_\rho \equiv 0$ modulo $\mathfrak{N}(\Gamma)^\rho$. Hence, for $\rho \leqq i$, the elements of $\mathfrak{N}(\Gamma)^\rho$ and only these elements are mapped upon elements of $\mathfrak{N}(\Gamma')^\rho$. This implies that $\mathfrak{K}_\rho(\mathfrak{G})$ is mapped upon $\mathfrak{K}_\rho(\mathfrak{G}')$ and hence $\mathfrak{K}_\rho(\mathfrak{G})/\mathfrak{H} \simeq \mathfrak{K}_\rho(\mathfrak{G}')$ for $\rho \leqq i$.

5. Consider any central series of $\mathfrak{G}$

$$\mathfrak{G} = \mathfrak{F}_1 \supseteq \mathfrak{F}_2 \supseteq \mathfrak{F}_3 \supseteq \cdots \tag{5.1}$$

such that

(1) $(\mathfrak{F}_i, \mathfrak{G}) \subseteq \mathfrak{F}_{i+1}$,

(2) $F_i^p \in \mathfrak{F}_{i_p}$ if $F_i \in \mathfrak{F}_i$.

The $\mathfrak{K}$-series is such a one, with the stronger property (2.4), (1).

Among all the series (5.1) there is a minimal series

$$\mathfrak{G} = \mathfrak{M}_1 \supseteq \mathfrak{M}_2 \supseteq \mathfrak{M}_3 \supseteq \cdots \tag{5.2}$$

which is defined by

$$\mathfrak{M}_i = \{(\mathfrak{M}_{i-1}, \mathfrak{G}), \mathfrak{M}_{(i/p)}^{(p)}\},$$

(i/p) being the least integer $\geqq i/p$, and $\mathfrak{M}_\lambda^{(p)}$ the set of all pth powers of elements of $\mathfrak{M}_\lambda$. It is easily verified that $\mathfrak{F}_i \supseteq \mathfrak{M}_i$, for any series (5.1), and in particular

$$\mathfrak{K}_i \supseteq \mathfrak{M}_i. \tag{5.3}$$

We call the series (5.2) the $\mathfrak{M}$-series of $\mathfrak{G}$. Where necessary we write $\mathfrak{M}_\lambda = \mathfrak{M}_\lambda(\mathfrak{G})$ to stress the fact that we are discussing the $\mathfrak{M}$-series of the particular group $\mathfrak{G}$. It is obvious that we have the following:

THEOREM 5.4. *If $\mathfrak{H}$ is a self-conjugate subgroup of $\mathfrak{G}$, then the subgroup $\mathfrak{M}_\lambda(\mathfrak{G})$ maps on the subgroup $\mathfrak{M}_\lambda(\mathfrak{G}/\mathfrak{H})$ in the homomorphism of $\mathfrak{G}$ upon $\mathfrak{G}/\mathfrak{H}$.*

We proceed to prove:

THEOREM 5.5[11]. *The $\mathfrak{K}$-series and $\mathfrak{M}$-series of $\mathfrak{G}$ are identical.*

---

[11] I am indebted to Professor R. Brauer for the definition of the $\mathfrak{M}$-series and for the subsequent proof of Theorem 5.5.

Proof. Suppose that we know already $\mathfrak{M}_i = \mathfrak{K}_i$ for $i = 1, 2, \cdots, t$, so that we have

$$\begin{aligned} \mathfrak{G} = \mathfrak{M}_1 = \mathfrak{K}_1 \supseteq \mathfrak{M}_2 = \mathfrak{K}_2 \supseteq \cdots \supseteq \mathfrak{M}_t = \mathfrak{K}_t \\ \supseteq \mathfrak{K}_{t+1} \supseteq \mathfrak{M}_{t+1} \supseteq \cdots . \end{aligned} \tag{5.6}$$

We apply (4.1) and (5.4) with $\mathfrak{H} = \mathfrak{M}_{t+1}$ and write $\mathfrak{G}' = \mathfrak{G}/\mathfrak{M}_{t+1}$. We then have

$$\begin{aligned} \mathfrak{G}' = \mathfrak{M}_1(\mathfrak{G}') = \mathfrak{K}_1(\mathfrak{G}') \supseteq \cdots \supseteq \mathfrak{M}_t(\mathfrak{G}') \\ = \mathfrak{K}_t(\mathfrak{G}') \supseteq \mathfrak{K}_{t+1}(\mathfrak{G}') \supseteq \mathfrak{M}_{t+1}(\mathfrak{G}') = \{1\}. \end{aligned} \tag{5.6'}$$

If we can prove $\mathfrak{K}_{t+1}(\mathfrak{G}') = \{1\}$, it will follow that $\mathfrak{K}_{t+1}(\mathfrak{G}) = \mathfrak{M}_{t+1}(\mathfrak{G})$, and this will finish our proof. If we replace $\mathfrak{G}'$ by $\mathfrak{G}$, we see that it is sufficient to treat the following case, namely, that in (5.6) $\mathfrak{M}_{t+1} = \{1\}$, in which we have to show that also $\mathfrak{K}_{t+1} = \{1\}$.

Let n_L be an element of $\mathfrak{N}^L$. As in Theorem 3.2, and because $\mathfrak{N}^{L+1} = 0$, n_L is a linear combination of products $\prod_i (F_{1,\rho_i} - 1)$, with L factors. As before, consider more generally all products

$$\pi = \prod (F_{\sigma_i,\tau_i} - 1) \tag{5.7}$$

with $\sum_i \sigma_i = L$, and $\sigma_i \leqq t$ (any order of factors admitted).

LEMMA 5.8. *Every product π of the type* (5.7) *may be expressed linearly by the products*

$$M_\lambda^{(L)} = \prod_{i,\nu} (F_{i,\nu} - 1)^{\beta_{i,\nu}}, \qquad \nu = 1, 2, \cdots, d_i;\ i = 1, 2, \cdots, t,$$

where $0 \leqq \beta_{i,\nu} < p$, and $\sum_{i,\nu} i\beta_{i,\nu} = L$, the factors being in the natural order as in (2.5).

To prove this lemma we use a method similar to that used in Theorem 3.2. A modification is necessary because we postulated only the property (5.1), (1) for the $\mathfrak{M}$-series, rather than the stronger $(\mathfrak{M}_i, \mathfrak{M}_j) \subseteq \mathfrak{M}_{i+j}$.

Suppose the lemma false and let π be the "highest" product which cannot be expressed linearly in terms of the $M_\lambda^{(L)}$. As before[12], we show that the same property is enjoyed by the product obtained by interchanging two consecutive factors $(A-1)$ and $(B-1)$ of π. Let $\pi = \pi_1(B-1)(A-1)\pi_2$ and suppose $A \in \mathfrak{K}_\rho$, $B \in \mathfrak{K}_\sigma$, $\rho, \sigma \leqq t$. Now $(A-1)$ may be replaced by an expression

$$(A-1) \equiv \sum c_{\lambda_1 \cdots \lambda_\rho} (F_{1,\lambda_1} - 1) \cdots (F_{1,\lambda_\rho} - 1) \text{ modulo } \mathfrak{N}^{\rho+1}, \tag{5.9}$$

$$c_{\lambda_1 \cdots \lambda_\rho} \in GF(p),$$

and since $\mathfrak{N}^{L+1} = 0$, and $\pi \in \mathfrak{N}^L$, we may replace $(A-1)$ by this sum in the expression for π above, and obtain

[12] Compare the proof of Theorem 3.2.

$$\pi = \sum c_{\lambda_1 \cdots \lambda_\rho} \pi_1 (B-1)(F_{1,\lambda_1}-1) \cdots (F_{1,\lambda_\rho}-1)\pi_2.$$

As before, we use the identity (3.5) to interchange successively $(B-1)$ and $(F_{1,\lambda_1}-1), \cdots, (F_{1,\lambda_\rho}-1)$. When, in one of the terms of π, $(B-1)(F_{1,\lambda_i}-1)$ is replaced by $(F_{1,\lambda_i}-1)(B-1)$, we get, in the expression for π, two additional terms T_1 and T_2. In the first, the two interchanged factors $(B-1)(F_{1,\lambda_i}-1)$ are replaced by $(F_{1,\lambda_i}B-1)((B, F_{1,\lambda_i})-1)$ and in the second by $((B, F_{1,\lambda_i})-1)$. The first term T_1 vanishes since it is in $\mathfrak{N}^{L+1}=0$. The other term T_2 is transformed as follows: (B, F_{1,λ_i}) lies in $(\mathfrak{K}_\sigma, \mathfrak{G}) = (\mathfrak{M}_\sigma, \mathfrak{G}) \subseteq \mathfrak{M}_{\sigma+1}$. If $\sigma = t$, then $((B, F_{1,\lambda_i})-1)=0$, since $\mathfrak{M}_{t+1}=1$; if $\sigma<t$, we have $\mathfrak{K}_{\sigma+1}=\mathfrak{M}_{\sigma+1}$. We replace, using (2.7), $((B, F_{1,\lambda_i})-1)$ by a linear combination, modulo $\mathfrak{N}^{\sigma+2}$ of terms $(F_{\sigma+1,\nu}-1)$. Hence this term T_2 becomes a sum of terms "higher" than π and of a term in $\mathfrak{N}^{L+1}$. The latter vanishes, the first terms can be expressed by the $M_\lambda^{(L)}$. This shows that T_2 can also be disregarded. Finally, $\pi_1(B-1)(A-1)\pi_2$ will be replaced by

$$\pi_1\Big(\sum c_{\lambda_1 \cdots \lambda_\rho}(F_{1,\lambda_1}-1)\cdots(F_{1,\lambda_\rho}-1)(B-1)\Big)\pi_2 = \pi_1(A-1)(B-1)\pi_2.$$

Since π could not be expressed linearly by the M_λ^L, this new product $\pi_1(A-1)(B-1)\pi_2$ cannot be expressed by them either.

We may therefore pick π in the form

$$\pi = \prod (F_{i,\nu}-1)^{\beta_{i,\nu}}, \qquad \sum_{i,\nu} i\beta_{i,\nu} = L.$$

If all $\beta_{i,\nu}<p$, then π is itself an $M_\lambda^{(L)}$, which is a contradiction. Suppose $\beta_{i,\nu} \geqq p$. Then we write $(F_{i,\nu}-1)^p = (F_{i,\nu}^p - 1)$: $F_{i,\nu} \in \mathfrak{K}_i = \mathfrak{M}_i$, and hence $F_{i,\nu}^p \in \mathfrak{M}_{ip}$. If $ip>t$, $\mathfrak{M}_{ip}=1$, and $F_{i,\nu}^p=1$, which would lead to $\pi=0$. If $ip \leqq t$, then $\mathfrak{M}_{ip}=\mathfrak{K}_{ip}$, and we replace $(F_{i,\nu}-1)^p$ by a combination of the $(F_{ip,\mu}-1)$, which gives rise to terms all "higher" than π. This again is a contradiction, and the lemma is proven.

The proof of Theorem 5.5 now follows at once. There is an $\mathfrak{M}_\lambda^L \neq 0$, since $\mathfrak{N}^L \neq 0$, and hence $L = \sum i\beta_{i,\nu} \leqq \sum i(p-1)d_i$, where $\nu = 1, 2, \cdots, d_i$; $i = 1, 2, \cdots, t$. However, $L = \sum j(p-1)d_j$, where $j = 1, 2, \cdots, e$ if $\mathfrak{K}_e \neq 1$, $\mathfrak{K}_{e+1}=1$. Comparing we see $t \geqq e$, and hence $\mathfrak{K}_{t+1}=\mathfrak{K}_{e+1}=1$, which proves $\mathfrak{M}_{t+1}=\mathfrak{K}_{t+1}$ as required.

COROLLARY 5.10. $(\mathfrak{M}_i, \mathfrak{M}_j) \subseteq \mathfrak{M}_{i+j}$.

6. We conclude with a discussion of the $\mathfrak{K}$-series of two special types of p-groups, namely, groups all of whose elements except the identity are of order p, and abelian groups. Lombardo-Radici[15] has investigated the structure of the radical of the group ring of an abelian group by elementary methods, and we shall show that the application of the more powerful ideas of the present paper leads to identical results.

[15] L. Lombardo-Radici, Rendiconti del Seminario Matematico della Università di Roma, (4), vol. 2 (1938), p. 312.

Consider first a group $\mathfrak{G}$ which contains no element of order greater than p. Clearly conditions (2.4), (2) and (5.1), (2), become trivially satisfied, and the $\mathfrak{K}$-series of $\mathfrak{G}$ is defined by the relation (5.2), (1), which becomes

$$\mathfrak{K}_1 = \mathfrak{G}, \qquad \mathfrak{K}_i = (\mathfrak{K}_{i-1}, \mathfrak{G}). \tag{6.1}$$

These relations, however, imply that the $\mathfrak{K}$-series is the "lower central series" of $\mathfrak{G}$(14). That is, $\mathfrak{K}_\lambda = \mathfrak{M}_\lambda = \mathfrak{H}_\lambda$, where $\mathfrak{G} = \mathfrak{H}_1 \supset \mathfrak{H}_2 \supset \cdots \supset \mathfrak{H}_e \supset 1$ is the lower central series of $\mathfrak{G}$.

THEOREM 6.2. *The $\mathfrak{K}$-series of a group which contains no elements whose order is greater than p is identical with the lower central series of the group. The "length" of the $\mathfrak{K}$-series of such a group is equal to the class of the group.*

Now let $\mathfrak{A}$ be an abelian group of type $(p^{\mu_1}, p^{\mu_2}, \cdots, p^{\mu_d})$, where $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_d$. In this case conditions (2.4), (1) and (5.1), (1), become trivially satisfied, and the $\mathfrak{K}$-series of $\mathfrak{A}$ is defined by the relation (5.2), (2), which becomes

$$\mathfrak{K}_1 = \mathfrak{A}, \qquad \mathfrak{K}_i = \{\mathfrak{K}^p_{(i/p)}\}. \tag{6.3}$$

Let $\mathfrak{A}^{(\lambda)}$ be the subgroup of the p^λth powers of elements of $\mathfrak{A}$. It is readily verified that

$$\mathfrak{K}_{p^\lambda} = \mathfrak{A}^{(\lambda)} \subset \mathfrak{K}_{p^\lambda+1} = \cdots = \mathfrak{K}_{p^{\lambda+1}} = \mathfrak{A}^{(\lambda+1)}, \qquad \lambda = 0, 1, 2, \cdots, \mu_1. \tag{6.4}$$

In particular, since $\mathfrak{A}^{(\mu_1-1)} \neq 1$, $\mathfrak{A}^{(\mu_1)} = 1$, we have

THEOREM 6.5. *The $\mathfrak{K}$-series of an abelian group of type $(p^{\mu_1}, p^{\mu_2}, \cdots, p^{\mu_d})$, $\mu_1 \geqq \cdots \geqq \mu_d$, has length p^{μ_1-1}.*

From Theorem 3.7, we deduce that for the group ring of an abelian group the rank of $\mathfrak{N}^w/\mathfrak{N}^{w+1}$ is equal to the coefficient of x^w in the expansion of

$$(1 + x + \cdots + x^{(p-1)})^{\delta_0}(1 + x^p + \cdots + x^{p(p-1)})^{\delta_1} \cdots (1 + x^{p^\lambda} + \cdots + x^{p^\lambda(p-1)})^{\delta_\lambda} \cdots,$$

where p^{δ_λ} is the order of $\mathfrak{A}^{(\lambda)}/\mathfrak{A}^{(\lambda+1)}$. By remarking that δ_λ is equal to the number of the integers μ_ρ which exceed λ, it is readily verified that the above product is equal to

$$(1 + x + \cdots + x^{p^{\mu_1}-1})(1 + x + \cdots + x^{p^{\mu_2}-1}) \cdots (1 + x + \cdots + x^{p^{\mu_d}-1}),$$

which is the form in which Lombardo-Radici obtained his result.

UNIVERSITY OF TORONTO,
TORONTO, ONTARIO, CANADA,
YALE UNIVERSITY,
NEW HAVEN, CONN.

(14) Cf. Hall, loc. cit., §2.

INEQUALITIES FOR HARMONIC POLYNOMIALS IN TWO AND THREE DIMENSIONS

BY

A. C. SCHAEFFER AND G. SZEGÖ

In what follows we discuss certain inequalities involving harmonic polynomials of two and of three variables, that is, polynomials $u(x, y)$ and $u(x, y, z)$ which satisfy Laplace's equation $u_{xx}+u_{yy}=0$ and $u_{xx}+u_{yy}+u_{zz}=0$, respectively. The inequalities in question furnish bounds for these polynomials and for their derivatives under proper conditions. Of particular interest are inequalities of the type of S. Bernstein's theorem, as discussed by the second author in a recent paper [7](1).

The first part of the present paper deals with the two-dimensional and the second part with the three-dimensional case. Of fundamental importance throughout the paper is an interpolation formula which is stated and proved in §2 of Part I. In fact the results of Part I may be regarded as systematic applications of this formula. Several inequalities of this part are generalizations and refinements of earlier theorems. Most of the problems of the second part are new.

In an Appendix we consider a generalization of the main problem treated in Part II, and another problem which deals with ellipses and is only in loose relationship with the other topics considered in the present paper.

PART I. TWO-DIMENSIONAL HARMONIC POLYNOMIALS

1. DEFINITIONS AND NOTATIONS

1. We consider polynomials $u(x, y)$ of degree n in the cartesian coordinates x, y satisfying Laplace's differential equation. Let $n \geqq 1$. Introducing polar coordinates r, ϕ we write $u(x, y) = U(r, \phi)$. The standard representation

$$U(r, \phi) = a_0 + 2\sum_{m=1}^{n} r^m(a_m \cos m\phi + b_m \sin m\phi) \tag{1}$$

is well known. For the sake of simplicity we assume (except in the Appendix, §2) that the a_m and b_m are real, although many of the following results can be extended to harmonic polynomials with complex coefficients. The boundary function $U(1, \phi) = U(\phi)$ is the most general real trigonometric polynomial of degree n,

$$U(\phi) = a_0 + 2\sum_{m=1}^{n} (a_m \cos m\phi + b_m \sin m\phi). \tag{1'}$$

Presented to the Society, April 6, 1940; received by the editors September 11, 1940.

(1) Numbers in brackets refer to the Bibliography at the end of the text.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

The conjugate harmonic polynomial $V(r, \phi)$ is unique except for an additive constant; we write

$$V(r, \phi) = 2\sum_{m=1}^{n} r^m(-b_m \cos m\phi + a_m \sin m\phi). \tag{2}$$

Then $U(r, \phi)$ and $V(r, \phi)$ are the real and imaginary parts, respectively, of the polynomial

$$F(z) = a_0 + 2\sum_{m=1}^{n} (a_m - ib_m)z^m \tag{3}$$

when $z = re^{i\phi}$.

2. Let K_n be the class of trigonometric polynomials $U(\phi)$ of degree n satisfying the condition

$$|U(\nu\pi/n)| \leqq 1, \qquad 0 \leqq \nu \leqq 2n - 1.$$

Replacing this condition by

$$U(\nu\pi/n) \geqq 0, \qquad 0 \leqq \nu \leqq 2n - 1,$$

we refer to the corresponding trigonometric polynomials of degree n as those of the class L_n.

3. Let $\lambda_0, \mu_0, \lambda_1, \mu_1, \cdots, \lambda_n, \mu_n$ be given real numbers with $\mu_0 = \mu_n = 0$. In the following development we are interested in obtaining inequalities for the expression

$$G = \lambda_0 a_0 + 2\sum_{m=1}^{n} (\lambda_m a_m + \mu_m b_m) \tag{4}$$

where a_m and b_m are the coefficients of an arbitrary trigonometric polynomial of the class K_n or L_n. In these investigations the trigonometric polynomial

$$\begin{aligned} g(\phi) &= 2\sum_{m=0}^{n}{}' (\lambda_{n-m} \cos m\phi - \mu_{n-m} \sin m\phi) \\ &= 2\sum_{m=0}^{n}{}' (\lambda_m \cos (n - m)\phi - \mu_m \sin (n - m)\phi) \end{aligned} \tag{5}$$

will be of primary importance. Here the symbol $\sum_{m=0}^{\prime n}$ means that the first and last terms ($m = 0$ and $m = n$) are to be multiplied by $\frac{1}{2}$. The above definitions and notations are used throughout the text. Of course (4) can be written in the form

$$G = \frac{1}{2\pi}\int_{-\pi}^{\pi} U(\phi)\left\{\lambda_0 + 2\sum_{m=1}^{n} (\lambda_m \cos m\phi + \mu_m \sin m\phi)\right\} d\phi.$$

2. Fundamental formulas and theorems

1. Our starting point is the following important

Interpolation formula. *Given the real numbers* $\lambda_0, \mu_0, \lambda_1, \mu_1, \cdots, \lambda_n, \mu_n$ *with* $\mu_0=\mu_n=0$ *and the trigonometric polynomial* $U(\phi)$ *of degree* n, *we have by using the notations* (1'), (4), (5)

$$G = \frac{1}{2n}\sum_{\nu=0}^{2n-1}(-1)^\nu g(\nu\pi/n)U(\nu\pi/n). \tag{6}$$

Special cases of (6) due to M. Riesz [2] and Szegö were discussed in [4]. These cases were used in the proof of S. Bernstein's theorem and the "gradient theorem" (cf. §4) of Szegö, respectively. The interpolation formula of M. Riesz [2] is the special case corresponding to $\lambda_m=\cos m\alpha$, $0\leqq m\leqq n$, $\mu_m=\sin m\alpha$, $0\leqq m\leqq n-1$; it is the following:

$$\begin{aligned} U(\alpha) &= 2b_n \sin n\alpha \\ &\quad + \frac{1}{2n}\sum_{\nu=0}^{2n-1} U(\nu\pi/n)\{1 + 2\cos(\alpha-\nu\pi/n) + 2\cos 2(\alpha-\nu\pi/n) + \cdots \\ &\quad + 2\cos(n-1)(\alpha-\nu\pi/n) + \cos n(\alpha-\nu\pi/n)\} \\ &= 2b_n \sin n\alpha + \frac{\sin n\alpha}{2n}\sum_{\nu=0}^{2n-1}(-1)^\nu U(\nu\pi/n)\cot(\alpha/2 - \nu\pi/(2n)). \end{aligned} \tag{7}$$

We also note the formula for the corresponding harmonic polynomials (regarding (8) see [5, p. 75]):

$$\begin{aligned} U(r,\alpha) &= 2b_n r^n \sin n\alpha + \frac{1}{2n}\sum_{\nu=0}^{2n-1} U(\nu\pi/n)\{1 + 2r\cos(\alpha-\nu\pi/n) \\ &\quad + 2r^2\cos 2(\alpha-\nu\pi/n) + \cdots \\ &\quad + 2r^{n-1}\cos(n-1)(\alpha-\nu\pi/n) + r^n\cos n(\alpha-\nu\pi/n)\}, \end{aligned} \tag{8}$$

$$\begin{aligned} V(r,\alpha) &= -2b_n r^n \cos n\alpha + \frac{1}{2n}\sum_{\nu=0}^{2n-1} U(\nu\pi/n)\{2r\sin(\alpha-\nu\pi/n) \\ &\quad + 2r^2\sin 2(\alpha-\nu\pi/n) + \cdots \\ &\quad + 2r^{n-1}\sin(n-1)(\alpha-\nu\pi/n) + r^n\sin n(\alpha-\nu\pi/n)\}. \end{aligned} \tag{9}$$

2. From the formula of M. Riesz, (6) can easily be derived. A direct proof of (6) can be based on the well known remark that for any trigonometric polynomial $f(\phi)$ of degree $2n-1$,

$$\frac{1}{2n}\sum_{\nu=0}^{2n-1} f(\nu\pi/n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\phi)d\phi. \tag{10}$$

Now,

$$(11)\qquad \begin{aligned}(-1)^{\nu}\cos[(n-m)\nu\pi/n] &= \cos(m\nu\pi/n),\\ (-1)^{\nu}\sin[(n-m)\nu\pi/n] &= -\sin(m\nu\pi/n),\end{aligned}$$

so making use of (5), the right-hand side of (6) is equal to

$$\frac{1}{2n}\sum_{\nu=0}^{2n-1}{\sum_{m=0}^{n}}'2U(\nu\pi/n)(\lambda_m\cos(m\nu\pi/n)+\mu_m\sin(m\nu\pi/n)).$$

On changing the order of summation and using (10) we find that this is equal to G. It has to be taken into account that

$$\frac{1}{2n}\sum_{\nu=0}^{2n-1}\lambda_n\cos(n\nu\pi/n)(2a_n\cos(n\nu\pi/n)+2b_n\sin(n\nu\pi/n))=2\lambda_n a_n.$$

3. As a first consequence of this interpolation formula we prove

THEOREM A. *Let $\lambda_0, \mu_0, \lambda_1, \mu_1, \cdots, \lambda_n, \mu_n$ be given real numbers, $\mu_0=\mu_n=0$, for which*

$$(12)\qquad g(\nu\pi/n)\geqq 0, \qquad 0\leqq\nu\leqq 2n-1.$$

If $U(\phi)$ belongs to the class K_n, then

$$(13)\qquad |G|=\left|\lambda_0 a_0+2\sum_{m=1}^{n}(a_m\lambda_m+\mu_m b_m)\right|\leqq\lambda_n.$$

If (12) is not satisfied (that is, if for at least one value of ν the sign $<$ holds), then there is a member of K_n for which $G>\lambda_n$.

REMARK. *Let $\nu_1, \nu_2, \cdots, \nu_l$ denote all the values of ν, $0\leqq\nu\leqq 2n-1$, for which in (12) the sign $>$ holds. Then $|G|=\lambda_n$ if and only if $(-1)^{\nu}U(\nu\pi/n)=\gamma$, $\nu=\nu_1, \nu_2, \cdots, \nu_l$. Here γ represents either $+1$ or -1. If in all the inequalities (12) the sign $>$ holds, then $|G|=\lambda_n$ if and only if $U(\phi)=\pm\cos n\phi+c\sin n\phi$, c real.*

This result was proved by Szegö in the less precise form that (13) follows if, in place of (12) and the condition that $U(\phi)$ belongs to K_n, we have $g(\phi)\geqq 0$ and $|U(\phi)|\leqq 1$ for all real ϕ. The above generalization of the theorem was proved by S. Bernstein [1] using the method of best approximation. The present method is similar to Szegö's.

4. Inequality (13) is a simple consequence of the fundamental interpolation formula; for under the hypotheses of Theorem A,

$$(14)\qquad |G|=\left|\frac{1}{2n}\sum_{\nu=0}^{2n-1}(-1)^{\nu}g(\nu\pi/n)U(\nu\pi/n)\right|\leqq\frac{1}{2n}\sum_{\nu=0}^{2n-1}g(\nu\pi/n)=\lambda_n.$$

Now let $g(k\pi/n)<0$ for a special value of k. We construct a trigonometric polynomial of class K_n for which $G>\lambda_n$. The function $f(\phi)$

$=(2n)^{-1}\sin n\phi\cot(\phi/2)$ is a trigonometric polynomial of degree n vanishing at $\phi=\nu\pi/n$, $1\leqq\nu\leqq 2n-1$, and equal to 1 at $\phi=0$. [This function serves as a basis for the interpolation formula of M. Riesz. See (7).] Then the trigonometric polynomial defined by

$$U(\phi)=\sum_{\nu=0}^{2n-1}(-1)^{\nu}f(\phi-\nu\pi/n)\operatorname{sgn}g(\nu\pi/n) \tag{15}$$

belongs to K_n. But from (6)

$$G=\frac{1}{2n}\sum_{\nu=0}^{2n-1}\left|g(\nu\pi/n)\right|>\frac{1}{2n}\sum_{\nu=0}^{2n-1}g(\nu\pi/n)=\lambda_n. \tag{16}$$

But more than this is true. If $g(k\pi/n)<0$ for a special k, there is a trigonometric polynomial of degree n which is bounded by 1 for all real ϕ, but for which $G>\lambda_n$. To show this let ϵ be a small positive number and let

$$U_\epsilon(\phi)=(1-\epsilon-\epsilon^{3/2})(\cos n\phi+\epsilon U(\phi))$$

where $U(\phi)$ is defined by (15). Suppose $U_\epsilon(\phi)$ attains its maximum at ϕ_1, which of course must lie within a distance $\pi/(2n)$ of some point $\nu\pi/n$. Then

$$\epsilon U'(\phi_1)=n\sin n\phi_1,$$

so, from the inequality $(\sin\theta)/\theta\geqq 2/\pi$, $|\theta|\leqq\pi/2$,

$$\epsilon\left|U'(\phi_1)\right|=n\left|\sin n(\phi_1-\nu\pi/n)\right|\geqq(2n/\pi)\left|\phi_1-\nu\pi/n\right|.$$

Thus ϕ_1 lies in fact within a distance $c_1\epsilon$ of some point $\nu\pi/n$, c_1 independent of ϵ. Then the mean value theorem furnishes

$$U_\epsilon(\phi_1)-U_\epsilon(\nu\pi/n)=(1-\epsilon-\epsilon^{3/2})(\phi_1-\nu\pi/n)(-n\sin n\theta+\epsilon U'(\theta))$$

where $\left|\theta-\nu\pi/n\right|\leqq\left|\phi_1-\nu\pi/n\right|\leqq c_1\epsilon$. The last factor on the right will be less than $c_2\epsilon$, c_2 independent of ϵ, so

$$\left|U_\epsilon(\phi_1)\right|<(1-\epsilon-\epsilon^{3/2})(1+\epsilon+c_1c_2\epsilon^2)<1$$

for all small ϵ.

But

$$G=G_\epsilon=(1-\epsilon-\epsilon^{3/2})/(2n)\sum_{\nu=0}^{2n-1}g(\nu\pi/n)\{1+\epsilon\operatorname{sgn}g(\nu\pi/n)\}$$

$$\geqq(1-\epsilon-\epsilon^{3/2})\{\lambda_n+\epsilon\lambda_n+(\epsilon/n)\left|g(k\pi/n)\right|\}>\lambda_n$$

if ϵ is small enough.

If all $g(\nu\pi/n)$ are greater than zero the equality occurs in (14) if and only if $(-1)^{\nu}U(\nu\pi/n)=\gamma$, $0\leqq\nu\leqq 2n-1$. Here γ represents $+1$ or -1. In this case, however, the function $\gamma U(\phi)-\cos n\phi$ vanishes at $\nu\pi/n$, $0\leqq\nu\leqq 2n-1$, and so is identically equal to $c\sin n\phi$ where c is a real constant. The more general statement in the Remark to Theorem A is also clear. The values of $U(\nu\pi/n)$ for the remaining $\nu\neq\nu_1,\nu_2,\cdots,\nu_l$ are obviously immaterial.

5. A result which is closely related to Theorem A is the following:

THEOREM B. *Let $\lambda_0,\mu_0,\lambda_1,\mu_1,\cdots,\lambda_n,\mu_n$ be given real numbers, $\mu_0=\mu_n=0$, for which*

$$(-1)^\nu g(\nu\pi/n) \geqq 0, \qquad 0 \leqq \nu \leqq 2n-1. \tag{17}$$

If $U(\phi)$ is a member of the class L_n, then

$$G = \lambda_0 a_0 + 2\sum_{m=1}^{n} (\lambda_m a_m + \mu_m b_m) \geqq 0. \tag{18}$$

If (17) is not satisfied (that is, if for at least one value of ν the sign $<$ holds), then there is a member of L_n for which $G<0$.

REMARK. *Let $\nu_1, \nu_2, \cdots, \nu_l$ denote all the values of ν, $0 \leqq \nu \leqq 2n-1$, for which in (17) the sign $>$ holds. Then $G=0$ if and only if $U(\nu\pi/n)=0$, $\nu=\nu_1, \nu_2, \cdots, \nu_l$. If in all the inequalities (17) the sign $>$ holds, then $G=0$ if and only if*

$$U(\phi) \equiv c \sin n\phi, \qquad c \text{ real.}$$

Inequality (18) follows immediately from (6). Now suppose $(-1)^k g(k\pi/n) < 0$ for a special value of k. Then

$$U(\phi) = (1/2n) \sin n(\phi - k\pi/n) \cot \tfrac{1}{2}(\phi - k\pi/n)$$

belongs to the class L_n, and $U(k\pi/n)=1$. But for this function

$$G = \frac{1}{2n} \sum_{\nu=0}^{2n-1} (-1)^\nu g(\nu\pi/n) U(\nu\pi/n) = \frac{1}{2n} (-1)^k g(k\pi/n) U(k\pi/n) < 0.$$

If for all ν, $(-1)^\nu g(\nu\pi/n) > 0$, the equality occurs in (18) if and only if $U(\nu\pi/n)=0$, $0 \leqq \nu \leqq 2n-1$. But this implies that $U(\phi) \equiv c \sin n\phi$, c real. The more general assertion of the Remark to Theorem B is also clear.

6. Theorems A and B are "dual" theorems in the following sense. Let $\lambda_0, \mu_0, \lambda_1, \mu_1, \cdots, \lambda_n, \mu_n; \mu_0=\mu_n=0$; be given real numbers, and suppose that the function $g_1(\phi) = 2\sum'_{m=0}{}^{n}(\lambda_{n-m} \cos m\phi - \mu_{n-m} \sin m\phi)$ satisfies (12). If $\{\lambda_m', \mu_m'\}$ is a new sequence defined by $\lambda_m' = \lambda_{n-m}$, $\mu_m' = -\mu_{n-m}$, the function $g_2(\phi) = 2\sum'_{m=0}{}^{n}(\lambda'_{n-m} \cos m\phi - \mu'_{n-m} \sin m\phi)$ will satisfy (17). In fact it follows from (11) that $g_2(\nu\pi/n) = (-1)^\nu g_1(\nu\pi/n)$. Thus if the sequence $\{\lambda_m, \mu_m\}$ is such that Theorem A is applicable to polynomials of the class K_n, the sequence $\{\lambda_m', \mu_m'\}$ will be such that Theorem B will be applicable to polynomials of the class L_n. The converse is also true.

7. Finally we note some identities which will be useful in the later considerations (cf. [5, p. 75]):

$$\begin{aligned} A &= 2\sum_{m=0}^{n}{}' r^m \cos m\phi \\ &= \frac{(1-r^2)(1-r^n \cos n\phi) + 2r^{n+1} \sin n\phi \sin \phi}{1 - 2r\cos\phi + r^2}, \end{aligned} \tag{19}$$

$$B = 2\sum_{m=0}^{n}{}' r^{n-m}\cos m\phi$$

(20)

$$= \frac{(r^2-1)(r^n-\cos n\phi)+2r\sin n\phi\sin\phi}{1-2r\cos\phi+r^2},$$

(21) $$C = 2\sum_{m=0}^{n}{}'\cos m\phi = \sin n\phi\cot(\phi/2).$$

For $\phi=\nu\pi/n$, ν an integer, we have

(22) $$\operatorname{sgn} A = +1, \qquad \operatorname{sgn} B = (-1)^r \quad \text{if} \quad r<1,$$

(23) $$\operatorname{sgn} A = (-1)^r, \qquad \operatorname{sgn} B = +1 \quad \text{if} \quad r>1,$$

and

(24) $$\operatorname{sgn} C = 0 \quad \text{or} \quad +1$$

according as ν is not or is divisible by $2n$.

3. Applications

1. In this paragraph we consider several applications of Theorems A and B.

I. *Let $U(r,\phi)$ be a harmonic polynomial and let $U(1,\phi)=U(\phi)$ belong to the class K_n. Then for $R>1$*

(25) $$|U(R,\nu\pi/n)+\delta R^n U(R^{-1},\nu\pi/n)| \leqq R^n+\delta, \qquad 0\leqq\nu\leqq 2n-1,$$

where $\delta=+1$ or -1. In the special case $\nu=\nu_0$ of (25) the equality occurs if and only if $U(1,\nu\pi/n)=\gamma$ either for all even ν or for all odd ν according as $\nu_0+\frac{1}{2}(1-\delta)$ is even or odd; here $\gamma=+1$ or -1.

From this follows

I′. *Let $U(r,\phi)$ be a harmonic polynomial of degree n satisfying the condition $|U(1,\phi)|\leqq 1$, ϕ arbitrary real. Then for $R>1$,*

(25′) $$|U(R,\phi)+\delta R^n U(R^{-1},\phi)| \leqq R^n+\delta, \qquad \phi \text{ arbitrary real},$$

where $\delta=+1$ or -1. The equality occurs in (25′) if and only if $\pm U(r,\phi)\equiv 1-\lambda+\lambda r^n\cos n(\phi-\phi_0)$ where ϕ_0 is real, $0\leqq\lambda\leqq 1$.

To prove I it will be sufficient to consider the special case $\nu=0$. Let

$$\lambda_m = R^m+\delta R^{n-m}, \quad \mu_m = 0, \qquad 0\leqq m\leqq n,$$

where $\delta=\pm 1$. Then

$$g(\phi) = 2\sum_{m=0}^{n}{}' R^{n-m}\cos m\phi + 2\delta\sum_{m=0}^{n}{}' R^{n-m}\cos(n-m)\phi,$$

and making use of (11) we obtain

$$g(\nu\pi/n) = 2(1 + \delta(-1)^{\nu}) \sum_{m=0}^{n}{}' R^{n-m} \cos(m\nu\pi/n). \tag{26}$$

The last sum is positive if $R>1$ [see (23)]. Consequently $g(\nu\pi/n)$ is greater than zero either for all even ν or for all odd ν, and vanishes for the remaining ν. Then (13) furnishes

$$\left| a_0(1 + \delta R^n) + 2\sum_{m=1}^{n} a_m(R^m + \delta R^{n-m}) \right| \leqq R^n + \delta,$$

that is, inequality (25). The conditions of equality in (25) follow from the Remark to Theorem A. Suppose for example that $\nu_0=0$ and $\delta=+1$. Then (26) shows that $g(\nu\pi/n)>0$ for even ν and $=0$ for odd ν. Thus the equality holds if and only if $U(1, \nu\pi/n)=\gamma$ for even ν, whereas $U(1, \nu\pi/n)$ is arbitrary for odd ν.

2. Inequality (25) is true if δ is not required to be ± 1, but is allowed to have any value between -1 and $+1$. This follows by the argument used in the proof of I, since (26) is non-negative if $-1 \leqq \delta \leqq 1$. It is interesting to note, however, that if $-1<\delta<1$ inequality (25) may become an equality, but under conditions different from those stated in I. For if $-1<\delta<1$, then $g(\nu\pi/n)>0$ for all ν, so (25) becomes an equality for the narrower class of functions $U(r, \phi) = \pm r^n \cos n\phi + cr^n \sin n\phi$, c real.

3. Inequality (25′) is obtained from (25) in a familiar way. However the question of (25′) becoming an equality requires special consideration. Let the sign "$=$" hold in (25′) for $\phi=0$, and let $\delta=+1$. In virtue of the result on the conditions of equality in I we have $U(1, \nu\pi/n)=\gamma$ for $\nu=0, 2, 4, \cdots, 2n-2$. Then the derivative of $U(1, \phi)$ with respect to ϕ must vanish at the same points. Thus, if λ is any real constant the function $f(\phi)=\gamma U(1, \phi)-\cos n\phi - \lambda(1-\cos n\phi)$ will have n double zeros, mod 2π. But choosing $\lambda=\gamma U(1, \pi/(2n))$ there will also be a zero at $\phi=\pi/2n$; that is, $f(\phi)$ will have $2n+1$ zeros, mod 2π, and so will vanish identically. This furnishes $\gamma U(1, \phi) \equiv \lambda + (1-\lambda) \cos n\phi$; in order that $|U(1, \phi)| \leqq 1$ be satisfied it is necessary that $0 \leqq \lambda \leqq 1$; finally we substitute $1-\lambda$ for λ.

4. The dual of I is

II. *Let $U(r, \phi)$ be a harmonic polynomial and let $U(1, \phi)=U(\phi)$ belong to the class L_n. Then for $R>1$*

$$|U(R, \nu\pi/n)| \leqq R^n U(R^{-1}, \nu\pi/n), \qquad 0 \leqq \nu \leqq 2n-1. \tag{27}$$

The equality holds if and only if $U(1, \nu\pi/n)=0$ either for all even ν or for all odd ν.

A consequence of this is the following theorem of Szegö [3, p. 333]:

II′. *Let $U(r, \phi)$ be a harmonic polynomial of degree n satisfying $U(1, \phi) \geqq 0$, ϕ arbitrary real. Then for $R>1$,*

$$|U(R, \phi)| \leqq R^n U(R^{-1}, \phi), \qquad \phi \text{ arbitrary real.} \tag{27′}$$

The equality occurs if and only if $U(r, \phi) \equiv \lambda\{1+r^n \cos n(\phi-\phi_0)\}$, ϕ_0 and λ real, $\lambda \geqq 0$.

To prove II it is again sufficient to consider the case $\nu=0$. Let $\lambda_m = R^{n-m} + \delta R^m$, $\mu_m = 0$ where, as before, $R>1$ and $\delta = \pm 1$. Then by the remark of §2.6 and by (26), the function $g(\phi)$ constructed with the present choice of λ_m and μ_m will satisfy (17). More precisely, $(-1)^\nu g(\nu\pi/n)$ will be greater than zero either for all even ν or for all odd ν, and will vanish for the remaining ν in the range $0 \leqq \nu \leqq 2n-1$. Thus, according to Theorem B

$$a_0(R^n + \delta) + 2\sum_{m=1}^{n} (R^{n-m} + \delta R^m) a_m \geqq 0$$

or

$$R^n U(R^{-1}, 0) + \delta U(R, 0) \geqq 0. \tag{28}$$

This shows that (27) is true for $\nu=0$.

If (27) is an equality (for $\nu=0$) then (28) will be also for proper choice of δ. Then the Remark to Theorem B shows that $U(1, \nu\pi/n)$ must vanish, either for all even ν or for all odd ν. Conversely, if $U(1, \nu\pi/n)=0$, ν even, we choose $\delta=+1$; then according to (26), $g(\nu\pi/n)=0$ for all odd ν so that in (28) the equality holds. The same is true in (27). The argument is similar if $U(1, \nu\pi/n)=0$ for all odd ν.

5. From the extensions of I and I′ given in §3.2 we obtain for $\delta=0$:

If $U(r, \phi)$ satisfies the conditions of I, *then for $R>1$*

$$|U(R, \nu\pi/n)| \leqq R^n, \qquad 0 \leqq \nu \leqq 2n-1. \tag{29}$$

If $U(R, \phi)$ satisfies the conditions of I′, *then for $R>1$*

$$|U(R, \phi)| \leqq R^n, \qquad \phi \text{ arbitrary real.} \tag{30}$$

The second of these is already known [3]. In the second part of the present paper we deal with the space analogues of (29) and (30).

6. Here we insert a direct proof of II′ different from that given above or in [3]. It is sufficient to prove (27′) for the special case $\phi=0$; also considering $\frac{1}{2}\{U(r, \phi)+U(r, -\phi)\}$ instead of $U(r, \phi)$ it suffices to prove (27′) for the case in which $\phi=0$ and $U(r, \phi)$ contains only cosine terms, that is, for

$$U(r, \phi) = a_0 + 2\sum_{m=1}^{n} a_m r^m \cos m\phi.$$

Finally we can assume $U(r, \phi)>0$ for $0 \leqq r<1$. Now $U(r, \phi)$ is the real part of the polynomial

$$F(z) = F(re^{i\phi}) = a_0 + 2\sum_{m=0}^{n} a_m z^m \neq 0, \qquad |z|<1,$$

so $p(z)=z^nF(z^{-1})/F(z)$ is regular for $|z|<1$. According to a familiar argument, $p(z)$ has the constant modulus 1 for $|z|=1$, so

$$|z^nF(z^{-1})|<F(z), \qquad |z|<1,$$

unless $p(z)\equiv\text{const.}=\pm 1$. This inequality furnishes (27′) by choosing $z=R^{-1}$, $R>1$. The assumption $p(z)\equiv\pm 1$ implies $a_0=\pm 2a_n$, from which, since $U(1,\phi)\geqq 0$, we conclude([5]) that $U(1,\phi)\equiv\lambda(1\pm\cos n\phi)$, $\lambda>0$.

Returning again to the general case, the equality can hold in (27′) at $\phi=0$ only if

$$U(1,\phi)=\lambda(1\pm\cos n\phi)+S(\phi)$$

where $S(\phi)$ is a sine polynomial. This again implies $U(1,\phi)\equiv\lambda(1\pm\cos n\phi)$.

7. An inequality of S. Bernstein states:

Let $U(\phi)$ be a trigonometric polynomial of degree n satisfying $|U(1,\phi)|\leqq 1$, ϕ arbitrary real. Then $|U'(\phi)|\leqq n$, ϕ arbitrary real, with the equality if and only if $U(\phi)\equiv\cos n(\phi-\phi_0)$, ϕ_0 real.

This will be a consequence of the following([6]):

Let $U(\phi)$ belong to the class K_n. Then $|U'(\pi/(2n))|\leqq n$ with the equality if and only if $U(\phi)\equiv\pm\cos n\phi+c\sin n\phi$, c real.

To prove this, let

$$\lambda_m=m\sin(m\pi/(2n)), \qquad \mu_m=-m\cos(m\pi/(2n)), \qquad 0\leqq m\leqq n.$$

Then

$$g(\phi)=2\sum_{m=0}^{n}{}'(n-m)\{\cos(m\pi/(2n))\cos m\phi+\sin(m\pi/(2n))\sin m\phi\}$$
$$=\left\{\frac{\sin\frac{1}{2}n(\pi/(2n)-\phi)}{\sin\frac{1}{2}(\pi/(2n)-\phi)}\right\}^2$$

is the classical Fejér kernel, and $g(\nu\pi/n)>0$. It follows from Theorem A that

$$\left|2\sum_{m=1}^{n}m\{a_m\sin(m\pi/(2n))-b_m\cos(m\pi/(2n))\}\right|\leqq\lambda_n$$

which is the assertion.

4. The "gradient theorem"

1. Szegö proved the following results [4] which for ease of reference we call the "gradient theorem" and the "generalized gradient theorem." They are refinements of S. Bernstein's theorem.

([5]) L. Fejér, *Über trigonometrische Polynome*, Journal für die reine und angewandte Mathematik, vol. 146 (1915), pp. 53–82; p. 72.

([6]) Cf. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, 1925, Problem IV, p. 201, and pp. 35, 218, 219; see in particular p. 219.

GRADIENT THEOREM. *If $U(r, \phi)$ is a harmonic polynomial of degree n satisfying $|U(r, \phi)| \leqq 1$ in $r \leqq 1$, then $|\text{grad } U| \leqq n$ in $r \leqq 1$. The equality holds if and only if $U(r, \phi) \equiv r^n \cos n(\phi - \phi_0)$, ϕ_0 real.*

GENERALIZED GRADIENT THEOREM. *If $U(r, \phi)$ is a harmonic polynomial of degree n satisfying $|U(r, \phi)| \leqq 1$ in $r \leqq 1$, then*

$$|\text{grad } U(r, \phi)| + n|\sigma_n(r, \phi)| \leqq n \qquad \text{in } r \leqq 1. \tag{31}$$

Except at points $(1, \phi)$ where $|U(1, \phi)| = 1$ the equality occurs if and only if $\pm U(r, \phi) \equiv 1 - \lambda + \lambda r^n \cos n(\phi - \phi_0)$, λ and ϕ_0 real, $0 \leqq \lambda \leqq 1$.

Here and henceforth $\sigma_n(r, \phi)$ denotes the nth Cesàro means of first order; if $U(r, \phi)$ is given by (1), then

$$\sigma_n(r, \phi) = a_0 + 2\sum_{m=1}^{n} (1 - m/n) r^m (a_m \cos m\phi + b_m \sin m\phi). \tag{32}$$

If $U(r, \phi)$ and $V(r, \phi)$ are conjugate harmonic polynomials we note that

$$\begin{aligned} |\text{grad } U| &= \{(r^{-1}U_\phi)^2 + (r^{-1}V_\phi)^2\}^{1/2} = \{(U_r)^2 + (r^{-1}U_\phi)^2\}^{1/2} \\ &= \{(U_r)^2 + (V_r)^2\}^{1/2}. \end{aligned} \tag{33}$$

In particular it follows from the gradient theorem that $|U_r| \leqq n$ and $|U_\phi| \leqq n$. The last inequality is S. Bernstein's theorem.

2. Incidentally the conditions of equality for the generalized gradient theorem were not discussed in [4]. To obtain this theorem by the method of the present paper we first show that

$$|\cos \alpha U_r(1, \alpha/n) + \sin \alpha V_r(1, \alpha/n) + n\sigma_n(1, \alpha/n)| \leqq n \tag{34}$$

for every real α. Now $U(r, \phi - \beta)$, β real, also satisfies the conditions of the generalized gradient theorem; so (34) will imply that

$$|\cos \alpha U_r(1, \phi) + \sin \alpha V_r(1, \phi) + n\sigma_n(1, \phi)| \leqq n.$$

Inequality (31) will be an immediate consequence of this, at least for $r = 1$.

To prove (34) let

$$\lambda_m = m \cos(\alpha - m\alpha/n) + (n - m) \cos(m\alpha/n),$$

$$\mu_m = -m \sin(\alpha - m\alpha/n) + (n - m) \sin(m\alpha/n).$$

Then

$$g(\phi) = 2\sum_{m=0}^{n}{}' (n-m) \cos(m\phi - m\alpha/n) + 2\sum_{m=0}^{n}{}' (n - m) \cos(m\alpha/n + n\phi - m\phi)$$

$$g(\nu\pi/n) = 2\{1 + (-1)^r\} \sum_{m=0}^{n}{}' (n - m) \cos (m(\alpha - \nu\pi)/n)$$

$$= \{1 + (-1)^r\} \left\{ \frac{\sin ((\alpha - \nu\pi)/2)}{\sin ((\alpha - \nu\pi)/(2n))} \right\}^2 \geqq 0.$$

A little algebraic manipulation will show that $|G|$ is equal to the left-hand side of (34), hence (34) is a consequence of Theorem A.

To discuss the conditions of equality note that $g(\nu\pi/n) = 0$ if and only if ν is odd or $\alpha - \nu\pi$ is of the form $\alpha - \nu\pi = 2k\pi$ where k is an integer not divisible by n. Consequently, if α/π is not an integer (34) is an equality if and only if $U(1, \nu\pi/n) = \gamma$, $\gamma = \pm 1$, for all even ν. If α/π is an integer the same is true unless $\alpha \equiv 0 \pmod{2\pi}$. In this exceptional case we see that the left-hand side of (34) is equal to

$$| U_r(1, \alpha/n) + n\sigma_n(1, \alpha/n) | = | nU(1, \alpha/n) |$$

and this is equal to n only at points where $U(1, \alpha/n) = \pm 1$.

The same argument which was used in §3.3 shows that at points on the unit circle where $|U(1, \phi)| < 1$ the equality holds only if $\pm U(1, \phi) \equiv 1 - \lambda + \lambda \cos n(\phi - \phi_0)$, $0 \leqq \lambda \leqq 1$, ϕ_0 real.

This proves the generalized gradient theorem for $r = 1$. If U_s denotes the directional derivative of U in a fixed direction s, then $U_s(r, \phi) + n\sigma_n(r, \phi)$ is a harmonic function. It follows that

$$| U_s(r, \phi) + n\sigma_n(r, \phi) | \leqq n, \qquad r < 1. \tag{35}$$

But now, if (r, ϕ) is a fixed point in the interior of the unit circle and s is allowed to vary we have

$$| \operatorname{grad} U(r, \phi) | + n | \sigma_n(r, \phi) | = \max_s | U_s(r, \phi) + n\sigma_n(r, \phi) | \leqq n.$$

According to the maximum principle, (35) can be an equality only if the corresponding equality is identically satisfied. In this case $U_s(1, \phi) + n\sigma_n(1, \phi) \equiv \pm n$ for all real ϕ so $\pm U(1, \phi) \equiv 1 - \lambda + \lambda \cos n(\phi - \phi_0)$.

3. In the further course of this paragraph we prove the following generalization of the gradient theorems.

III. *Let $U(r, \phi)$ be a harmonic polynomial and let $U(1, \phi) = U(\phi)$ belong to the class K_n. Then if $p > q \geqq 0$, $p + q \geqq 2$, and α is real,*

$$\begin{aligned} &| \cos \alpha \{ U(p, \alpha/n) - U(q, \alpha/n) \} + \sin \alpha \{ V(p, \alpha/n) - V(q, \alpha/n) \} | \\ &\quad + | p^n U(p^{-1}, \alpha/n) - q^n U(q^{-1}, \alpha/n) | \leqq p^n - q^n. \end{aligned} \tag{36}$$

The equality occurs if and only if $U(1, \nu\pi/n) = \gamma$, $\gamma = \pm 1$, either for all even ν or for all odd ν, the only exception being the case $n = 2$, $p + q = 2$, α/π an integer.

From this follows:

III′. *Let $U(r, \phi)$ be a harmonic polynomial of degree n satisfying $|U(1, \phi)| \leqq 1$, ϕ arbitrary real. Then if $p > q \geqq 0$, $p+q \geqq 2$,*

$$[\{U(p, \phi) - U(q, \phi)\}^2 + \{V(p, \phi) - V(q, \phi)\}^2]^{1/2} + |p^n U(p^{-1}, \phi) - q^n U(q^{-1}, \phi)| \leqq p^n - q^n. \tag{36′}$$

The equality occurs if and only if $\pm U(r, \phi) \equiv \lambda + (1-\lambda) r^n \cos n(\phi - \phi_0)$, λ and ϕ_0 real, $0 \leqq \lambda \leqq 1$, the only exception being the case $n=2$, $p+q=2$(4).

These results include the generalized gradient theorem as a limiting case. For, let each side of (36) and (36′) be divided by $p-q$ and let $p \to 1$, $q \to 1$. We obtain

$$|\cos \alpha U_r(1, \alpha/n) + \sin \alpha V_r(1, \alpha/n)| + n|\sigma_n(1, \alpha/n)| \leqq n$$

and

$$[\{U_r(1, \phi)\}^2 + \{V_r(1, \phi)\}^2]^{1/2} + n|\sigma_n(1, \phi)| \leqq n,$$

respectively. The first inequality states slightly more than (34); the second inequality is the generalized gradient theorem.

4. To prove III let

$$\lambda_m = (p^m - q^m) \cos(\alpha - m\alpha/n) + \delta(p^{n-m} - q^{n-m}) \cos(m\alpha/n),$$
$$\mu_m = -(p^m - q^m) \sin(\alpha - m\alpha/n) + \delta(p^{n-m} - q^{n-m}) \sin(m\alpha/n).$$

Here δ represents $+1$ or -1. Then

$$g(\phi) = \sum_{m=0}^{n}{}' (p^{n-m} - q^{n-m}) \{\cos(m\alpha/n - m\phi) + \delta \cos(m\alpha/n + n\phi - m\phi)\}$$

so

$$g(\nu\pi/n) = \{1 + (-1)^\nu \delta\} \sum_{m=0}^{n}{}' (p^{n-m} - q^{n-m}) \cos(m\alpha/n - m\nu\pi/n).$$

If we can show that

$$F(\phi) = \sum_{m=0}^{n}{}' (p^{n-m} - q^{n-m}) \cos m\phi = \sum_{k=1}^{n} [\{(p-1)^k - (q-1)^k\} \sum_{m=0}^{n}{}' C_{n-m,k} \cos m\phi] \geqq 0, \quad \phi \text{ real}, \tag{37}$$

it will follow that $g(\phi)$ satisfies (12). But $\sum'_{m=0} C_{n-m,k} \cos m\phi \geqq 0$ if $k \geqq 1$, this expression being, apart from a positive factor, the Cesàro means of kth order of the series $\frac{1}{2} + \cos\phi + \cos 2\phi + \cdots$. Moreover, from the hypotheses of III

(4) In this case (36′) is an equality for $\phi = \phi'$ whenever $U(1, \phi') = \pm 1$. (Compare the generalized gradient theorem.)

we have $p-1>q-1$ and $p-1\geqq 1-q$, so $p-1\geqq|q-1|$. This furnishes (37). In case $n\geqq 3$, $F(\phi)$ is even *positive* since $(p-1)^3-(q-1)^3>0$ and the Cesàro means mentioned are *positive* if $k\geqq 2$. In this case $g(\nu\pi/n)>0$ either for all even ν or for all odd ν.

Now,

$$G = 2\sum_{m=1}^{n}(p^m - q^m)(a_m\cos(\alpha - m\alpha/n) - b_m\sin(\alpha - m\alpha/n))$$
$$+ \delta\{(p^n - q^n)a_0 + 2\sum_{m=1}^{n}(p^{n-m} - q^{n-m})(a_m\cos(m\alpha/n) + b_m\sin(m\alpha/n))\}$$
$$= \cos\alpha\{U(p, \alpha/n) - U(q, \alpha/n)\} + \sin\alpha\{V(p, \alpha/n) - V(q, \alpha/n)\}$$
$$+ \delta\{p^nU(p^{-1}, \alpha/n) - q^nU(q^{-1}, \alpha/n)\};$$

so an application of Theorem A proves (36).

5. If $n\geqq 3$ and the equality occurs in (36), G will be equal to $\pm\lambda_n$ for proper choice of δ. But, since $g(\nu\pi/n)>0$ either for all even ν or for all odd ν, we must have $U(1, \nu\pi/n)=\gamma$ for the same ν, where γ represents either $+1$ or -1.

The inverse statement is also clear. Assuming $U(1, \nu\pi/n)=\gamma$ for even ν we choose $\delta=+1$. Then (for $n\geqq 3$) $g(\nu\pi/n)=0$, ν odd, so $G=\gamma\lambda_n$ holds. This means that the equality holds in (36) since

$$|G| \leqq |\cos\alpha\{U(p, \alpha/n) - U(q, \alpha/n)\} + \sin\alpha\{V(p, \alpha/n) - V(q, \alpha/n)\}|$$
$$+ |p^nU(p^{-1}, \alpha/n) - q^nU(q^{-1}, \alpha/n)| \leqq \lambda_n.$$

In case $n=1$, $F(\phi)=\frac{1}{2}(p-q)>0$ so the same conditions of equality hold. This is true also for $n=2$ since $F(\phi)=(p-q)\{\frac{1}{2}(p+q-2)+2\cos^2(\phi/2)\}$, the only exception being the case $p+q=2$ when at the same time α is an integral multiple of π. [For the calculation of $g(\nu\pi/2)$ we need only $F(\frac{1}{2}(\alpha-\nu\pi))$, which is 0 only if $p+q=2$ and $(\alpha-\nu\pi)/2$ is an odd multiple of π.]

Regarding the special case $p>1$, $q=1$, see the joint paper of Rogosinski and Szegö in Mathematische Zeitschrift, vol. 28 (1928), pp. 73–94, in particular p. 81, (39).

6. The dual of III is:

IV. *Let $U(r, \phi)$ be a harmonic polynomial and let $U(1, \phi)=U(\phi)$ belong to the class L_n. Then if $p>q\geqq 0$, $p+q\geqq 2$,*

$$|\cos\alpha\{U(p, \alpha/n) - U(q, \alpha/n)\} + \sin\alpha\{V(p, \alpha/n) - V(q, \alpha/n)\}|$$
$$\leqq p^nU(p^{-1}, \alpha/n) - q^nU(q^{-1}, \alpha/n). \tag{38}$$

The equality occurs if and only if $U(1, \nu\pi/n)$ vanishes either for all even ν or for all odd ν, the only exception being the case $n=2$, $p+q=2$, α/π an integer.

From this follows:

IV′. *Let $U(r, \phi)$ be a harmonic polynomial of degree n satisfying $U(1, \phi) \geqq 0$, ϕ arbitrary real. Then if $p > q \geqq 0$ and $p+q \geqq 2$,*

$$[\{U(p, \phi) - U(q, \phi)\}^2 + \{V(p, \phi) - V(q, \phi)\}^2]^{1/2} \leqq p^n U(p^{-1}, \phi) - q^n U(q^{-1}, \phi). \tag{38′}$$

The equality occurs if and only if $U(r, \phi) \equiv \lambda\{1 + r^n \cos n(\phi - \phi_0)\}$, λ and ϕ_0 real, $\lambda \geqq 0$, the only exception being the case $n = 2$, $p + q = 2$.

To prove IV we set

$$\lambda_m = (p^{n-m} - q^{n-m}) \cos (m\alpha/n) + \delta(p^m - q^m) \cos (\alpha - m\alpha/n),$$
$$\mu_m = (p^{n-m} - q^{n-m}) \sin (m\alpha/n) - \delta(p^m - q^m) \sin (\alpha - m\alpha/n)$$

where as before, $\delta = \pm 1$. In virtue of the remark of §2.6, it is only necessary to compute the corresponding quantity G, that is,

$$G = (p^n - q^n)a_0 + 2\sum_{m=1}^{n} (p^{n-m} - q^{n-m})(a_m \cos (m\alpha/n) + b_m \sin (m\alpha/n)) + 2\delta\sum_{m=1}^{n} (p^m - q^m)(a_m \cos (\alpha - m\alpha/n) - b_m \sin (\alpha - m\alpha/n)).$$

By Theorem B, we have $G \geqq 0$; or

$$p^n U(p^{-1}, \alpha/n) - q^n U(q^{-1}, \alpha/n) + \delta \cos \alpha\{U(p, \alpha/n) - U(q, \alpha/n)\} + \delta \sin \alpha\{V(p, \alpha/n) - V(q, \alpha/n)\} \geqq 0.$$

7. The dual of the generalized gradient theorem is the following:

V. *Let $U(r, \phi)$ be a harmonic polynomial of degree n satisfying $U(1, \phi) \geqq 0$. Then for $r \leqq 1$,*

$$|\operatorname{grad} U(r, \phi)| \leqq n\sigma_n(r, \phi). \tag{39}$$

The equality occurs if and only if $U(r, \phi) \equiv \lambda\{1 + r^n \cos n(\phi - \phi_0)\}$, λ and ϕ_0 real, $\lambda \geqq 0$. The only exception is the case in which $U(1, \phi_0) = 0$; then the equality holds in (39) for $r = 1$, $\phi = \phi_0$.

Incidentally V may be obtained by dividing each side of (38′) by $p - q$ and letting $p \to 1$, $q \to 1$. The left-hand side approaches $[\{U_r(1, \phi)\}^2 + \{V_r(1, \phi)\}^2]^{1/2}$, while the right-hand side approaches $n\sigma_n(1, \phi)$. This shows that (39) is true for $r = 1$. But then, by a familiar application of the maximum principle it follows for $r < 1$.

8. In terms of rational polynomials, III′ may be stated in the following equivalent form:

VI. *Let $F(z) = \sum_{m=0}^{n} c_m z^m$ be a polynomial of degree n satisfying $|\Re F(z)| \leqq 1$*

in $|z| \leqq 1$. *If* p *and* q *are real numbers,* $p > q \geqq 0$, $p+q \geqq 2$, *then for* $|z| \leqq 1$,

$$|F(pz) - F(qz)| + |\Re\{p^n F(p^{-1}z) - q^n F(q^{-1}z)\}| \leqq p^n - q^n. \tag{40}$$

The equality holds if and only if $F(z) \equiv az^n + b$ *where* $|a| + |\Re b| = 1$, *the only exception being the case* $n=2$, $p+q=2$[5].

Finally we observe that from the fact that (37) is non-negative the following refinement of a theorem of Rogosinski-Szegö follows, cf. loc. cit., p. 77, (21):

Let $\{s_n(z)\}$ *be the partial sums of a power series* $f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$ *convergent in the unit circle* $|z| < 1$ *and satisfying the condition* $|f(z)| \leqq 1$ *for* $|z| < 1$. *Then*

$$\left| \frac{r_1^{-n} s_n(r_1 z) - r_2^{-n} s_n(r_2 z)}{r_1^{-n} - r_2^{-n}} \right| \leqq 1$$

where $|z| \leqq 1$, $r_1 > 0$, $r_2 > 0$, $r_1 \neq r_2$, $r_1^{-1} + r_2^{-1} \geqq 2$.

9. Let

$$F(z) = \sum_{m=0}^{n} c_m z^m$$

be a polynomial of degree n satisfying the condition $|\Re F(z)| \leqq 1$ in $|z| \leqq 1$; according to the gradient theorem

$$|F'(z)| = \left| \sum_{m=1}^{n} m c_m z^{m-1} \right| \leqq n, \qquad |z| \leqq 1.$$

By repeated application of this theorem we obtain

$$\left| F^{(k)}(z) \right| \leqq \frac{n!}{(n-k)!}$$

or:

$$\left| \sum_{m=k}^{n} C_{m,k} c_m z^{m-k} \right| \leqq C_{n,k}; \qquad |z| \leqq 1, \ 1 \leqq k \leqq n. \tag{41}$$

In all these inequalities the sign $<$ holds unless $F(z) \equiv \epsilon z^n$, $|\epsilon| = 1$.

Using (41) we now prove:

VII. *Let* $\gamma_0, \gamma_1, \cdots, \gamma_n$ *be real numbers satisfying*

$$\delta_\nu = \sum_{m=0}^{\nu} (-1)^{\nu+m} C_{\nu,m} \gamma_m \geqq 0, \qquad 0 \leqq \nu \leqq n. \tag{42}$$

If $F(z) = \sum_{m=0}^{n} c_m z^m$ *is a polynomial of degree* n *with real or complex coefficients*

[5] See the footnote to III′.

satisfying $|\Re F(z)| \leqq 1$ *in* $|z| \leqq 1$ *we have*

$$\left| \sum_{m=0}^{n} \gamma_m c_m z^m \right| \leqq \gamma_n - \delta_0(1 - |F(z)|), \qquad |z| \leqq 1. \tag{43}$$

The equality holds if and only if $F(z) \equiv \epsilon z^n$, $|\epsilon| = 1$, $|z| = 1$, *or* $\gamma_0 = \gamma_1 = \gamma_2 = \cdots = \gamma_n = \gamma$.

If δ_ν is defined by (42) the known inversion formula

$$\gamma_m = \sum_{\nu=0}^{m} C_{m,\nu}\delta_\nu \tag{44}$$

may be proved in the following way. Expanding e^z and using Cauchy multiplication we verify that

$$\sum_{m=0}^{\infty} (-1)^m \frac{\delta_m}{m!} z^m = e^z \sum_{m=0}^{\infty} (-1)^m \frac{\gamma_m}{m!} z^m.$$

Bringing e^z to the left and again using Cauchy multiplication we obtain (44). Then (44) implies that

$$\sum_{m=0}^{n} \gamma_m c_m z^m = \sum_{\nu=0}^{n} \delta_\nu z^\nu \left\{ \sum_{m=\nu}^{n} C_{m,\nu} c_m z^{m-\nu} \right\};$$

so, according to (41),

$$\left| \sum_{m=0}^{n} \gamma_m c_m z^m \right| \leqq \delta_0 |F(z)| + \sum_{\nu=1}^{n} \delta_\nu |z|^\nu C_{n,\nu} \leqq \delta_0 |F(z)| + \gamma_n - \delta_0$$

in $|z| \leqq 1$. If the equality holds, then $F(z) \equiv \epsilon z^n$, $|\epsilon| = 1$, $|z| = 1$; or $\delta_1 = \delta_2 = \cdots = \delta_n = 0$. The second possibility is equivalent to the relation $\gamma_0 = \gamma_1 = \cdots = \gamma_n = \gamma$.

10. As a very simple application, (30) follows. Indeed, assuming that $U(r, \phi)$ is even, and defining $F(z)$ as in §3.6, the coefficients c_m are real. Putting $\gamma_m = R^m$, we obtain

$$\delta_\nu = \sum_{m=0}^{\nu} (-1)^{\nu-m} R^m C_{\nu,m} = (R-1)^\nu > 0.$$

Also $F(R) = U(R, 0)$, so from (43) for $z = 1$,

$$|U(R, 0)| = \left| \sum_{m=0}^{n} c_m R^m \right| \leqq R^n - (1 - |U(1, 0)|). \tag{45}$$

This is slightly sharper than (30).

11. Another simple application follows by choosing $\gamma_m = p^m - q^m$ where $p > q \geqq 0$. Then $\delta_\nu = (p-1)^\nu - (q-1)^\nu$ so (42) is equivalent with $p + q \geqq 2$ pro-

vided $n \geqq 2$. If $F(z)$ satisfies the conditions of VII it follows that

$$|F(p) - F(q)| \leqq p^n - q^n.$$

This is slightly less general than VI.

12. Finally we point out that the following classical theorem of A. Markoff (cf. [2, pp. 359–360]) is also a simple consequence of the gradient theorem:

Let $h(x)$ be a polynomial of degree n with real coefficients(6) *satisfying the condition $|h(x)| \leqq 1$ in $-1 \leqq x \leqq 1$. Then $|h'(x)| < n^2$ in $-1 < x < 1$, $n > 1$. For $n = 1$ we have $|h'(x)| \leqq 1$.*

Let

$$h(\cos\phi) = \sum_{m=0}^{n} \alpha_m \cos m\phi, \qquad F(z) = \sum_{m=0}^{n} \alpha_m z^m.$$

We have for $0 < \phi < \pi$, $z = e^{i\phi}$,

$$h'(\cos\phi) = \sum_{m=1}^{n} m\alpha_m \frac{\sin m\phi}{\sin\phi} = \sum_{m=1}^{n} m\alpha_m (z^m - \bar{z}^m)/(z - \bar{z})$$

$$= \frac{1}{z - \bar{z}} \int_{\bar{z}}^{z} \left\{ \sum_{m=1}^{n} m^2 \alpha_m \zeta^{m-1} \right\} d\zeta,$$

the integration extending along the chord from $\bar{z}$ to z. Since $|\Re F(z)| \leqq 1$, $|z| \leqq 1$, we have $|F'(z)| \leqq n$ and $|\{zF'(z)\}'| \leqq n^2$, $|z| \leqq 1$, so for $n > 1$, $|h'(\cos\phi)| < n^2$ follows.

5. Direct proof of V

1. Writing $U(\phi)$ and $V(\phi)$ for $U(1, \phi)$ and $V(1, \phi)$, respectively, to prove V it will be sufficient to show that

$$|U'(0) + iV'(0)| \leqq n\sigma_n(0)$$

where $\sigma_n(0)$ is the nth Cesàro means of first order of $U(0)$. The proof of this inequality can be based on a theorem of Fejér which states: *Let $U(\phi)$ be a trigonometric polynomial of degree n which is non-negative for all real values of ϕ. Then there exists a polynomial $A(z) = \sum_{\nu=0}^{n} a_\nu z^\nu$ such that $U(\phi) = |A(e^{i\phi})|^2$.*

In our case

$$U(\phi) = |A(e^{i\phi})|^2 = \left| \sum_{\nu=0}^{n} a_\nu e^{i\nu\phi} \right|^2$$

$$= \sum a_\nu \bar{a}_\mu \cos |\nu - \mu| \phi + i \sum a_\nu \bar{a}_\mu \operatorname{sgn}(\nu - \mu) \sin |\nu - \mu| \phi,$$

$$V(\phi) = \sum a_\nu \bar{a}_\mu \sin |\nu - \mu| \phi - i \sum a_\nu \bar{a}_\mu \operatorname{sgn}(\nu - \mu) \cos |\nu - \mu| \phi$$

and

$$U'(0) + iV'(0) = i \sum (\nu - \mu) a_\nu \bar{a}_\mu + i \sum |\nu - \mu| a_\nu \bar{a}_\mu, \tag{46}$$

$$n\sigma_n(0) = \sum (n - |\nu - \mu|) a_\nu \bar{a}_\mu. \tag{47}$$

In these and the following expressions the summation extends over the range $0 \leqq \nu \leqq n$, $0 \leqq \mu \leqq n$.

(6) This is not an essential restriction.

The quadratic form with coefficients $|\nu-\mu|$ was studied by Szegö [6]. He showed that the second sum in (46) may be written in the form

$$\sum |\nu-\mu|\, a_\nu \bar{a}_\mu = (n/2)\left|\sum_{\nu=0}^{n} a_\nu\right|^2 - (2/n)\sum_{m=1}^{n}\frac{\left|\sum_{\nu=0}^{n} a_\nu z_m^\nu\right|^2}{|1-z_m|^2} \tag{48}$$

where $z_m = e^{i(2m-1)\pi/n}$ is a root of the equation $z^n+1=0$. Thus the linear transformation

$$\begin{aligned} A_0 &= (n/2)^{1/2}(a_0 + a_1 + \cdots + a_n) = (n/2)^{1/2}A(1), \\ A_m &= (2/n)^{1/2}(a_0 + a_1 z_m + \cdots + a_n z_m^n)/(1-z_m) = (2/n)^{1/2}A(z_m)/(1-z_m), \quad 1 \leqq m \leqq n, \end{aligned} \tag{49}$$

furnishes the canonical representation

$$\sum |\nu-\mu|\, a_\nu \bar{a}_\mu = |A_0|^2 - \sum_{m=1}^{n} |A_m|^2.$$

Also, from (47)

$$n\sigma_n(0) = 2|A_0|^2 - |A_0|^2 + \sum_{m=1}^{n}|A_m|^2 = \sum_{m=0}^{n}|A_m|^2.$$

Next we express the first sum in (46) in terms of the A_m. Since

$$\sum \nu a_\nu \bar{a}_\mu = \left(\sum_{\mu=0}^{n}\bar{a}_\mu\right)\left(\sum_{\nu=0}^{n}\nu a_\nu\right) = \overline{A(1)}A'(1)$$

and since $\sum \mu a_\nu \bar{a}_\mu$ is the conjugate of this,

$$U'(0) + iV'(0) = -2\Im\{\overline{A(1)}A'(1)\} + i\left\{|A_0|^2 - \sum_{m=1}^{n}|A_m|^2\right\}. \tag{50}$$

Let $\omega(z) = (z^n+1)(z-1)$. Making use of (49), the Lagrange interpolation formula furnishes

$$A(z) = (n/2)^{1/2}\sum_{m=1}^{n}\frac{\omega(z)A_m(1-z_m)}{(z-z_m)\omega'(z_m)} + (2/n)^{1/2}A_0\frac{\omega(z)}{(z-1)\omega'(1)},$$

and this may be written

$$A(z) = (2n)^{-1/2}\sum_{m=1}^{n} A_m z_m \omega(z)/(z-z_m) + (2n)^{-1/2}A_0(z^n+1) \tag{51}$$

since $\omega'(z_m) = n(1-z_m)/z_m$, $\omega'(1)=2$. Differentiating (51) and setting $z=1$,

$$A'(1) = (2/n)^{1/2}\sum_{m=1}^{n} A_m z_m/(1-z_m) + (n/2)^{1/2}A_0. \tag{52}$$

Using Cauchy's inequality

$$\begin{aligned} |-2\Im\{\overline{A(1)}A'(1)\}|^2 &\leqq \left|(4/n)\bar{A}_0\sum_{m=1}^{n} A_m z_m/(1-z_m)\right|^2 \\ &\leqq (16/n^2)|A_0|^2\left\{\sum_{m=1}^{n}|A_m|^2\right\}\left\{\sum_{m=1}^{n}|1-z_m|^{-2}\right\} = 4|A_0|^2\sum_{m=1}^{n}|A_m|^2. \end{aligned} \tag{53}$$

In the last equality the special case $a_0=1$, $a_1=a_2=\cdots=a_n=0$, of (48) is used in order to obtain $\sum_{m=1}^{n}|1-z_m|^{-2}$. From this

$$|U'(0)+iV'(0)| \leqq \left\{4|A_0|^2\sum_{m=1}^{n}|A_m|^2 + \left(|A_0|^2 - \sum_{m=1}^{n}|A_m|^2\right)^2\right\}^{1/2} = |A_0|^2 + \sum_{m=1}^{n}|A_m|^2 \tag{54}$$

follows, which is the statement. The equality holds if and only if (53) is an equality; that is, if and only if either $A_0=0$, or $A_m = \lambda/(1-z_m)$, $m=1, 2, \cdots, n$, λ a constant and $\overline{A(1)}\{A'(1) - \frac{1}{2}nA(1)\}$ pure imaginary. The second condition implies $A(z_m) = \text{const.}$, $m=1, 2, \cdots, n$, so $A(z) = pz^n+q$ where p and q are constants. The additional condition regard-

ing $A(1)$ and $A'(1)$ furnishes $|p|^2=|q|^2$, so $A(z)=p(z^n+\gamma)$, $|\gamma|=1$. The other condition for equality, namely $A_0=0$, is equivalent with $U(0)=U(1, 0)=0$. This establishes the statement.

2. The generalized gradient theorem is an immediate consequence of V. If $U(r, \phi)$ satisfies the conditions of that theorem, then $1+\delta U(r, \phi)$ is non-negative in the unit circle where $\delta=+1$ or -1. Thus by what we have just shown

$$n+\delta n\sigma_n(r, \phi) \geqq |\operatorname{grad} U(r, \phi)|. \tag{55}$$

Here $1+\delta\sigma_n(r, \phi)$ is the Cesàro means of the harmonic polynomial $1+\delta U(r, \phi)$.

Part II. Three-dimensional harmonic polynomials

Introduction

In this part we deal with harmonic polynomials $u(x, y, z)$, that is, polynomials in the cartesian coordinates x, y, z satisfying Laplace's differential equation $\Delta u=0$. We frequently prefer to use polar coordinates r, θ, ϕ (in the usual notation) writing: $u(x, y, z)=U(r, \theta, \phi)$. Then the standard representation

$$\begin{aligned} U(r, \theta, \phi) = \sum_{\mu=0}^{n} r^{\mu} \Big\{ & a_{\mu 0} P_{\mu}(\cos\theta) \\ & + \sum_{\nu=1}^{\mu} \sin^{\nu}\theta P_{\mu}^{(\nu)}(\cos\theta)(a_{\mu\nu}\cos\nu\phi + b_{\mu\nu}\sin\nu\phi) \Big\} \end{aligned} \tag{1}$$

holds; here $P_n(x)$ is the nth polynomial of Legendre in the customary notation.

1. As already mentioned, our main task in Part II is to prove the three-dimensional analogue of inequality (30) of Part I, that is:

Let $u(x, y, z)$ be an arbitrary harmonic polynomial of degree n satisfying the condition

$$|U(x, y, z)| \leqq 1, \qquad x^2+y^2+z^2 \leqq 1. \tag{2}$$

Let $n \geqq 1$. Then at a point (x, y, z) in the exterior of the unit sphere, that is, for $x^2+y^2+z^2=R^2>1$,

$$\begin{aligned} |u(x, y, z)| \leqq \frac{1}{2} \frac{2\cdot 4\cdots 2n}{3\cdot 5\cdots(2n+1)} \Big\{ & (2n+1)R^n \\ & + (2n-3)\frac{n^2-(n+1)^2}{n^2-(n-2)^2} R^{n-2} \\ & + (2n-7)\frac{n^2-(n+1)^2}{n^2-(n-2)^2}\cdot\frac{n^2-(n-1)^2}{n^2-(n-4)^2} R^{n-4} \\ & + \cdots \Big\} = c_n(R). \end{aligned} \tag{3}$$

This bound $c_n(R)$ is attained exclusively for the polynomials $u(x, y, z)$ whose boundary values on the unit sphere appear in the form $\pm\cos n\gamma$ where γ is the spherical distance of the variable point $(1, \theta, \phi)$ from a fixed point $(1, \theta_0, \phi_0)$.

For the polynomials mentioned the bound $c_n(R)$ is attained only when (x, y, z) lies on the line connecting $(1, \theta_0, \phi_0)$ with the origin.

A direct characterization of $c_n(R)$ is the following: Write $\cos n\theta$ as a linear combination of Legendre functions $P_\nu(\cos\theta)$:

$$\cos n\theta = h_{n0}P_0(\cos\theta) + h_{n1}P_1(\cos\theta) + \cdots + h_{nn}P_n(\cos\theta). \tag{4}$$

Then

$$c_n(R) = h_{n0} + h_{n1}R + \cdots + h_{nn}R^n, \tag{5}$$

so $c_n(R)$ is a polynomial of degree n in R.

It is easy to show (§7) that for $R>1$

$$c_n(R) \simeq \tfrac{1}{2}\pi^{1/2}(1 - R^{-2})^{1/2}n^{1/2}R^n \quad \text{as} \quad n \to \infty, \tag{6}$$

so the result is slightly different in character from the corresponding result in the two-dimensional case in which the maximum in question is precisely R^n.

The proof of the Theorem mentioned requires a rather elaborate apparatus, in particular a discussion of the trigonometric polynomials

$$c_n(R) + 2\sum_{m=1}^{n-1} c_{n-m}(R)\cos m\theta + c_0(R)\cos n\theta. \tag{7}$$

This discussion is made possible by the following remarkable representation of $c_n(R)$:

$$c_n(R) = R^n + \int_{-1\le t\le 0,\,1\le t\le R} \frac{(R^2-1)(t^2-1)}{4t^2} \cdot \left\{\frac{t}{(R-t)(Rt-1)}\right\}^{3/2} (R^n - t^n)\,dt, \qquad R > 1. \tag{8}$$

We prove (8) by function-theoretic considerations applied to the "generating function" of the sequence $\{c_n(R)\}$.

2. In order to prove (3), say for $\theta=0$, we might try to use an argument similar to that in §4.10 of Part I. Since

$$U(R, \theta, \phi) = \sum_{m=0}^{n} \frac{(R-1)^m}{m!}\left\{\frac{\partial^m U(r, \theta, \phi)}{\partial r^m}\right\}_{r=1},$$

it would suffice to show that all the expressions in the braces $\{\ \}$ assume their maximum value if $U(1, \theta, \phi)=\cos n\theta$. This is indeed the case for $m=0$, and, according to a recent result of Szegö [7], also for $m=1$. In §6 of Part II

we shall prove this for $m=2$; however the proof (or refutation) of the general assertion seems to be difficult.

The theorem of Szegö just referred to states even more generally that under condition (2) the directional derivatives $|\partial u/\partial l|$ where $(x^2+y^2+z^2)^{1/2}=r=1$ and l is an arbitrary direction, assume their maximum value only for functions with boundary values of the form $\pm\cos n\gamma$. [This is of course equivalent to an estimate of $|\text{grad } u|$.] More generally, we might ask for the maximum of

$$\left|\frac{\partial}{\partial l_1}\frac{\partial}{\partial l_2}\cdots\frac{\partial}{\partial l_m}u\right|, \qquad r=1,$$

where $l_1, l_2, \cdots, l_m$ denote arbitrary directions. The former problem is a special case corresponding to coinciding directions $l_1, l_2, \cdots, l_m$.

1. Preliminaries

1. Considering the set of harmonic polynomials of degree n which satisfy condition (2), we see that the maximum of $|u|$ at an exterior point (x, y, z) is the same at all points on a sphere with center at the origin. Therefore we assume that $x=y=0$, $z=R$. Also, by taking the mean-value

$$(2\pi)^{-1}\int_{-\pi}^{\pi} U(r, \theta, \phi)d\phi,$$

we see that the maximum remains the same by restricting the set in question to the polynomials $u(x, y, z)=U(r, \theta, \phi)$ independent of ϕ. This leads to the following equivalent formulation of the problem:

Let $u(\theta)$ be a cosine polynomial of degree n,

$$u(\theta) = \sum_{m=0}^{n} \alpha_m P_m(\cos\theta) \tag{9}$$

expanded in terms of Legendre polynomials, and let $|u(\theta)| \leqq 1$ for $0 \leqq \theta \leqq \pi$. What is the maximum of $|\sum_{m=0}^{n}\alpha_m R^m|$?

2. Since

$$\alpha_m = \tfrac{1}{2}(2m+1)\int_0^{\pi} u(\theta)P_m(\cos\theta)\sin\theta\, d\theta, \tag{10}$$

we have

$$\sum_{m=0}^{n}\alpha_m R^m = \int_0^{\pi} u(\theta)\left\{\sum_{m=0}^{n}\tfrac{1}{2}(2m+1)P_m(\cos\theta)R^m\right\}\sin\theta\, d\theta. \tag{11}$$

We shall prove that the maximum in question is attained when $u(\theta)\equiv\pm\cos n\theta$

and that this maximum is

$$c_n(R) = \int_0^\pi \cos n\theta \left\{ \sum_{m=0}^{n} \tfrac{1}{2}(2m+1) P_m(\cos\theta) R^m \right\} \sin\theta \, d\theta. \tag{12}$$

This will imply (5). Representation (3) is obtained by taking into account the well known expression of $\cos n\theta$ in terms of Legendre polynomials(7).

3. The polynomial $c_n(R)$ contains only terms R^n, R^{n-2}, R^{n-4}, $\cdots$, the first coefficient being positive and all the remaining coefficients being negative. Since $c_n(1)=1$, this implies, as is easily seen, that all derivatives of $c_n(R)$ are positive when $R \geqq 1$. We have

$$\begin{aligned} &c_0(R) = 1, \quad c_1(R) = R, \quad c_2(R) = \tfrac{1}{3}R^2 - \tfrac{1}{3}, \\ &c_3(R) = \tfrac{2}{3}R^3 - \tfrac{3}{5}R, \quad c_4(R) = \tfrac{8}{35}R^4 - \tfrac{16}{21}R^2 - \tfrac{1}{15}. \end{aligned} \tag{13}$$

2. Generating function of the sequence $\{c_n(R)\}$

1. The following important result is to be proved.

Let $R>1$. The generating function of the sequence $\{c_n(R)\}$ is, for $|z|<R^{-1}$,

$$\begin{aligned} F(R,z) = F(z) &= \tfrac{1}{2}c_0(R) + \sum_{n=1}^{\infty} c_n(R) z^n \\ &= \tfrac{1}{2}(R^2-1) \frac{1-z^2}{\{(R-z)(1-Rz)\}^{3/2}} z^{1/2} \arctan\left(\frac{z(R-z)}{1-Rz}\right)^{1/2} \\ &\quad + \tfrac{1}{2}R \frac{1-z^2}{(R-z)(1-Rz)}. \end{aligned} \tag{14}$$

The determination of the multi-valued functions involved in this formula is evident: for small positive z we take $\{(R-z)(1-Rz)\}^{3/2}>0$, $z^{1/2}>0$, $\{z(R-z)/(1-Rz)\}^{1/2}>0$, and the branch of arc tan ω which vanishes at $\omega=0$.

The starting point of the proof is definition (12) of $c_n(R)$. We make use of Legendre functions of the second kind (cf. Szegö [7, p. 58]).

2. Let $Q_n(\xi)$ be Legendre functions of the second kind in the usual notation. It is well known that

$$\lim_{\epsilon \to +0} \{Q_n(u+i\epsilon) - Q_n(u-i\epsilon)\} = -i\pi P_n(u), \quad -1 < u < 1, \tag{15}$$

so if $z=re^{i\theta}$, $0<r<1$, $0<\theta<\pi$, then

$$P_n(\cos\theta) = \frac{2}{\pi} \lim_{r \to 1-0} \Im Q_n(\tfrac{1}{2}(z+z^{-1})). \tag{16}$$

---

(7) See E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, Cambridge University Press, 1931; p. 47, (48).

Hence

$$\sin\theta\sum_{m=0}^{n}\tfrac{1}{2}(2m+1)P_m(\cos\theta)R^m$$

$$=\lim_{r\to 1-0}\frac{z-z^{-1}}{2i}\,\frac{2}{\pi}\,\Im\left\{\sum_{m=0}^{n}\tfrac{1}{2}(2m+1)Q_m(\tfrac{1}{2}(z+z^{-1}))R^m\right\}$$

$$=\lim_{r\to 1-0}\Re\left\{\frac{1-z^2}{\pi z}\sum_{m=0}^{n}\tfrac{1}{2}(2m+1)Q_m(\tfrac{1}{2}(z+z^{-1}))R^m\right\}. \tag{17}$$

Now

$$Q_m(\tfrac{1}{2}(z+z^{-1}))$$

can be expanded into a power series of z convergent for $|z|<1$; the first non-vanishing term of this series is cz^{m+1}, $c\neq 0$. Also (8)

$$|Q_m(\tfrac{1}{2}(z+z^{-1}))|^{1/m}\simeq|z|, \qquad |z|<1,\ m\to\infty, \tag{18}$$

so the series

$$\sum_{m=0}^{\infty}\tfrac{1}{2}(2m+1)Q_m(\tfrac{1}{2}(z+z^{-1}))R^m$$

is convergent for $|z|\cdot R<1$.

According to (12) and (17), $c_n(R)$ is the real part of

$$\lim_{r\to 1-0}\frac{2}{\pi}\int_0^{\pi}\cos n\theta\cdot\left\{\frac{1-z^2}{2z}\sum_{m=0}^{n}\tfrac{1}{2}(2m+1)Q_m(\tfrac{1}{2}(z+z^{-1}))R^m\right\}d\theta.$$

This is simply the coefficient of z^n in the expansion of the function of z in the curled brackets. [For $n=0$ this gives twice the constant term of this expansion.] However, this coefficient does not change if the sum $\sum_{m=0}^{n}$ is replaced by $\sum_{m=0}^{\infty}$; so we obtain for the generating function (14) the representation

$$F(R;z)=\frac{1-z^2}{2z}\sum_{m=0}^{\infty}\tfrac{1}{2}(2m+1)Q_m(\tfrac{1}{2}(z+z^{-1}))R^m. \tag{19}$$

3. In order to calculate the last expression, we use the formula(9)

$$Q_m(\tfrac{1}{2}(z+z^{-1}))=\frac{1}{2}\int_{-\infty}^{+\infty}\{\tfrac{1}{2}(z+z^{-1})+\tfrac{1}{2}(z^{-1}-z)\cosh\tau\}^{-m-1}d\tau. \tag{20}$$

For sake of simplicity we assume that z is real and positive, $0<z<R^{-1}$. Since

$$\tfrac{1}{2}(z+z^{-1})+\tfrac{1}{2}(z^{-1}-z)\cosh\tau\geqq\tfrac{1}{2}(z+z^{-1})+\tfrac{1}{2}(z^{-1}-z)=z^{-1}>R,$$

(8) Cf. for instance, G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, 1939, p. 195, (8.23.2).

(9) See, for instance, Szegö, loc. cit., p. 90, (4.81.2).

term-by-term integration is permitted and we obtain

$$F(R;z) = \frac{1-z^2}{8z}\int_{-\infty}^{+\infty}\sum_{m=0}^{\infty}(2m+1)\left\{\tfrac{1}{2}(z+z^{-1})+\tfrac{1}{2}(z^{-1}-z)\cosh\tau\right\}^{-m-1}R^m d\tau$$

$$= \frac{1-z^2}{8z}\int_{-\infty}^{+\infty}\frac{\tfrac{1}{2}(z+z^{-1})+R+\tfrac{1}{2}(z^{-1}-z)\cosh\tau}{\left\{\tfrac{1}{2}(z+z^{-1})-R+\tfrac{1}{2}(z^{-1}-z)\cosh\tau\right\}^2}d\tau \tag{21}$$

$$= \frac{1}{4}\int_{-\infty}^{+\infty}\frac{A+\cosh\tau}{(B+\cosh\tau)^2}d\tau = \int_1^{\infty}\frac{1+t^2+2At}{(1+t^2+2Bt)^2}dt,$$

where

$$A = \frac{1+2Rz+z^2}{1-z^2}, \qquad B = \frac{1-2Rz+z^2}{1-z^2}. \tag{22}$$

Obviously $A>1$, $-1<B<+1$.

4. The primitive function of the last integrand in (21) is

$$\frac{1-AB}{(1-B^2)^{3/2}}\arctan\left\{\frac{t+B}{(1-B^2)^{1/2}}\right\}+\frac{B-A}{1-B^2}\,\frac{Bt+1}{1+t^2+2Bt}.$$

Hence, omitting trivial details, we find that

$$F(R;z) = \frac{(R^2-1)4z^2(1-z^2)}{\{4z(R-z)(1-Rz)\}^{3/2}}\left[\frac{\pi}{2}-\arctan\left(\frac{1}{z}\,\frac{1-Rz}{R-z}\right)^{1/2}\right] + \frac{2Rz(1-z^2)}{4z(R-z)(1-Rz)},$$

which easily furnishes (14).

Another way of writing (14) is

$$F(R;z) = \frac{1}{4i}(R^2-1)\frac{1-z^2}{\{(R-z)(1-Rz)\}^{3/2}}z^{1/2}\log\frac{1+i\left(\frac{z(R-z)}{1-Rz}\right)^{1/2}}{1-i\left(\frac{z(R-z)}{1-Rz}\right)^{1/2}}$$

$$+\tfrac{1}{2}R\frac{1-z^2}{(R-z)(1-Rz)} \tag{14'}$$

$$= G(R;z)+\tfrac{1}{2}R\frac{1-z^2}{(R-z)(1-Rz)}.$$

Obviously

$$G(R;z) = G(z) = \sum_{n=1}^{\infty}\left\{c_n(R)-\tfrac{1}{2}(R^n+R^{-n})\right\}z^n. \tag{23}$$

3. Discussion of the function $G(z)$

The function $G(R; z)=G(z)$ defined by (14′) and (23) is regular except perhaps at the points

$$-1,\ 0,\ R^{-1},\ 1,\ R,\ \infty.$$

We intend to discuss the behavior of $G(z)$ near these points; also, we wish to calculate the limits of $G(z)$ on the "upper" and "lower" border of the real segments $[-\infty, -1]$ and $[R^{-1}, 1]$.

1. $-\infty<z<-1$. First, for $-1<z<0$ we have

$$G(z)=\tfrac{1}{4}(R^2-1)\frac{1-z^2}{\{(R-z)(1-Rz)\}^{3/2}}(-z)^{1/2}\log\frac{1-\left(\dfrac{-z(R-z)}{1-Rz}\right)^{1/2}}{1+\left(\dfrac{-z(R-z)}{1-Rz}\right)^{1/2}}.$$

Here $\{\ \}^{3/2}>0$, $(-z)^{1/2}>0$, $(-z(R-z)/(1-Rz))^{1/2}>0$, and the log is real. This easily furnishes, for $-\infty<z<-1$,

$$(24)\quad G(z)=\tfrac{1}{4}(R^2-1)\frac{1-z^2}{\{(R-z)(1-Rz)\}^{3/2}}(-z)^{1/2}\left\{\log\frac{\left(\dfrac{-z(R-z)}{1-Rz}\right)^{1/2}-1}{\left(\dfrac{-z(R-z)}{1-Rz}\right)^{1/2}+1}\pm i\pi\right\}$$

where the signs $+$ and $-$ refer to the upper and lower border, respectively.

2. $R^{-1}<z<1$. First, for $0<z<R^{-1}$ we have the formula (14′) in which $\{\ \}^{3/2}>0$, $z^{1/2}>0$, $(z(R-z)/(1-Rz))^{1/2}>0$, and the log vanishes for $z=0$. If z increases from 0 to R^{-1}, the expression $z(R-z)/(1-Rz)$ increases from 0 to $+\infty$ and the imaginary part of the log increases from 0 to π.

If z describes a small circle around R^{-1}, from $(1-Rz)^{3/2}$, $z<R^{-1}$, we obtain by analytic continuation $(Rz-1)^{3/2}e^{\mp 3i\pi/2}$, $z>R^{-1}$, on the upper and lower border, respectively. Also, from

$$i\left(\frac{z(R-z)}{1-Rz}\right)^{1/2},\qquad z<R^{-1},$$

we obtain $\mp(z(R-z)/(Rz-1))^{1/2}$, $z>R^{-1}$. Thus we find for $z>R^{-1}$

$$G(z)=\frac{1}{4i}(R^2-1)e^{\pm 3i\pi/2}\frac{1-z^2}{\{(R-z)(Rz-1)\}^{3/2}}z^{1/2}\log\frac{1\mp\left(\dfrac{z(R-z)}{Rz-1}\right)^{1/2}}{1\pm\left(\dfrac{z(R-z)}{Rz-1}\right)^{1/2}}$$

where the imaginary part of the log approaches π as $z\to R^{-1}+0$.

This can be written in the form

$$(25)\quad G(z)=\tfrac{1}{2}(R^2-1)\frac{1-z^2}{\{(R-z)(Rz-1)\}^{3/2}}z^{1/2}\left\{\log\frac{1+\left(\dfrac{Rz-1}{z(R-z)}\right)^{1/2}}{1-\left(\dfrac{Rz-1}{z(R-z)}\right)^{1/2}}\mp i\pi\right\}$$

on the upper and lower border, respectively.

3. $1<z<R$. If z increases from R^{-1} to 1, the expression $(Rz-1)/(z(R-z))$ increases from 0 to 1; at $z=1$ the log becomes singular. Encircling $z=1$, we obtain

$$(26)\quad \begin{aligned} G(z) &= \tfrac{1}{2}(R^2-1)\frac{1-z^2}{\{(R-z)(Rz-1)\}^{3/2}}z^{1/2}\log\frac{1+\left(\dfrac{Rz-1}{z(R-z)}\right)^{1/2}}{\left(\dfrac{Rz-1}{z(R-z)}\right)^{1/2}-1} \\ &= \tfrac{1}{2}(R^2-1)\frac{1-z^2}{\{(R-z)(Rz-1)\}^{3/2}}z^{1/2}\log\frac{1+\left(\dfrac{z(R-z)}{Rz-1}\right)^{1/2}}{1-\left(\dfrac{z(R-z)}{Rz-1}\right)^{1/2}} \end{aligned}$$

on both borders. Here the log is real.

This shows that the branch of the function $G(z)$ which we have considered is single valued for $z>1$; it has a simple pole at $z=R$.

4. Integral representation of $c_n(R)$

1. Since $G(0)=0$, we have for $n\geqq 1$

$$(27)\quad c_n(R)-\tfrac{1}{2}(R^n+R^{-n})=\frac{1}{2\pi i}\int G(z)(z^{-n}-R^n)z^{-1}dz$$

where the integration curve consists of the following parts:

(a) a "large" circle $|z|=P$, $|\arc z|\leqq\pi-\epsilon$, $\epsilon\to+0$;

(b) the segment $[-P, -1]$ described twice, the upper border in the increasing, the lower in the decreasing way;

(c) the segment $[R^{-1}, 1]$ described twice, the upper border in the increasing, the lower in the decreasing way;

(d) a "small" circle around $z=R$ described in the negative sense.

The circle in (d) furnishes the negative residue corresponding to $z=R$.

2. Since

$$G(z)=O(|z|^{-1/2})\text{ as }z\to\infty,$$

the contribution of the large circle in (a) vanishes as the radius P approaches ∞.

According to (24) we obtain as the contribution of $[-\infty, -1]$:

$$\frac{1}{4}(R^2-1)\int_{-\infty}^{-1}\frac{1-z^2}{\{(R-z)(1-Rz)\}^{3/2}}(-z)^{1/2}(z^{-n}-R^n)z^{-1}dz$$

(28)

$$=-\frac{1}{4}(R^2-1)\int_{0}^{-1}\frac{1-t^{-2}}{\{(R-t^{-1})(1-Rt^{-1})\}^{3/2}}(-t)^{-1/2}(t^n-R^n)t^{-1}dt.$$

The contribution of $[R^{-1}, 1]$ is, according to (25),

$$-\frac{1}{4}(R^2-1)\int_{R^{-1}}^{1}\frac{1-z^2}{\{(R-z)(Rz-1)\}^{3/2}}z^{1/2}(z^{-n}-R^n)z^{-1}dz$$

(29)

$$=\frac{1}{4}(R^2-1)\int_{R}^{1}\frac{1-t^{-2}}{\{(R-t^{-1})(Rt^{-1}-1)\}^{3/2}}t^{-1/2}(t^n-R^n)t^{-1}dt.$$

These integrals are convergent.

Around $z=R$, the function $G(z)$ is single valued, and from (26) we find

$$G(z)=\frac{1}{4}(R^2-1)\frac{1-z^2}{\{(R-z)(Rz-1)\}^{3/2}}z^{1/2}\cdot 2\left(\frac{z(R-z)}{Rz-1}\right)^{1/2}+\cdots$$

$$=\frac{1}{2}(R^2-1)\frac{(1-z^2)z}{(Rz-1)^2}\frac{1}{R-z}+\cdots=\frac{1}{2}R(z-R)^{-1}+\cdots$$

where the terms not written out are regular at $z=R$. Thus the contribution of (d) is

$$-\frac{1}{2}R(R^{-n}-R^n)R^{-1}=\frac{1}{2}(R^n-R^{-n}).$$

Recapitulating, we obtain the important result

$$c_n(R)=R^n+\int_{-1\le t\le 0,\,1\le t\le R}p(R;t)(R^n-t^n)dt, \tag{30}$$

where

$$p(R;t)=p(t)=\frac{(R^2-1)(t^2-1)}{4t^2}\left\{\frac{t}{(R-t)(Rt-1)}\right\}^{3/2}. \tag{30'}$$

Formula (30) is trivial for $n=0$. We note that (except $t=0$) $\operatorname{sgn} p(R;t)=\operatorname{sgn}(t^2-1)$.

5. Solution of the maximum problem

1. Returning again to the maximum problem formulated in §1 of this part, we start with the remark that, as a generalization of (12),

$$c_\nu(R)=\int_0^\pi\cos\nu\theta\left\{\sum_{m=0}^{n}\frac{1}{2}(2m+1)P_m(\cos\theta)R^m\right\}\sin\theta\,d\theta,\qquad 0\le\nu\le n, \tag{31}$$

holds. Indeed,

$$\int_0^\pi \cos \nu\theta P_m(\cos\theta) \sin\theta \, d\theta = 0, \qquad \nu < m.$$

Therefore, the cosine expansion of the function $\pi\sum_{m=0}^{n}\frac{1}{2}(2m+1)P_m(\cos\theta)R^m \sin\theta$ begins with the terms

$$c_0(R) + 2\sum_{\nu=1}^{n} c_\nu(R) \cos \nu\theta.$$

Thus we obtain from (11)

$$\sum_{m=0}^{n} \alpha_m R^m = \pi^{-1} \int_0^\pi u(\theta) \left\{ c_0(R) + 2\sum_{\nu=1}^{n} c_\nu(R) \cos \nu\theta \right\} d\theta. \tag{32}$$

2. Application of Theorem A of Part I to the last expression shows that if

$$C_{\nu n}(R) = \sum_{m=0}^{n}{}' c_{n-m}(R) \cos(m\nu\pi/n) > 0, \qquad 0 \leqq \nu \leqq n, \tag{33}$$

then the maximum of (32) will be attained only by $u(\theta) = \pm\cos n\theta$. Here, as in Part I, the symbol $\sum'_{m=0}$ means that the terms $m=0$ and $m=n$ have to be multiplied by $\frac{1}{2}$. It is needless to say that $R>1$.

Our next purpose is to prove these inequalities. Note that having done this we have proved much more than the original statement, namely the inequality $\left|\sum_{m=0}^{n}\alpha_m R^m\right| \leqq c_n(R)$ under the condition that $|u(\nu\pi/n)| \leqq 1$, $0 \leqq \nu \leqq n$.

Since $c_n(R) > 0$ (cf. §1.3), it suffices to prove (33) for $1 \leqq \nu \leqq n$.

3. We combine the representation (30) with formula (20) of Part I and find that

$$\begin{aligned} C_{\nu n}(R) &= \tfrac{1}{2}(R^2-1)\frac{R^n + (-1)^{\nu-1}}{1 - 2R\cos(\nu\pi/n) + R^2} \\ &\quad + \int p(R;t)\left\{ \tfrac{1}{2}(R^2-1)\frac{R^n + (-1)^{\nu-1}}{1 - 2R\cos(\nu\pi/n) + R^2} \right. \\ &\quad \left. - \tfrac{1}{2}(t^2-1)\frac{t^n + (-1)^{\nu-1}}{1 - 2t\cos(\nu\pi/n) + t^2} \right\} dt \end{aligned} \tag{34}$$

where the integration extends over $-1 \leqq t \leqq 0$ and $1 \leqq t \leqq R$. Then to prove (33) it will be sufficient to show that

$$1 + \int p(R;t)\left\{ 1 - \frac{t^2-1}{R^2-1}\,\frac{t^n+(-1)^{\nu-1}}{R^n+(-1)^{\nu-1}}\,\frac{1-2R\cos(\nu\pi/n)+R^2}{1-2t\cos(\nu\pi/n)+t^2} \right\} dt > 0. \tag{35}$$

We consider only the case $n \geqq 3$, since (33) may be proved directly for $n=1, 2$. Now, let us consider the expression

$$-\{t^n+(-1)^{\nu-1}\}\frac{1-2R\lambda+R^2}{1-2t\lambda+t^2}=k(\lambda)$$

as a function of λ, $-1\leqq\lambda\leqq 1$. We have

$$k'(\lambda)=2(R-t)\{t^n+(-1)^{\nu-1}\}\frac{1-Rt}{(1-2t\lambda+t^2)^2}. \tag{36}$$

Let ν be *even*. In this case $k'(\lambda)<0$ whether $-1<t<0$ or $1<t<R$, so min $k(\lambda)$ is attained for $\lambda=+1$. Since (35) is positive for $\nu=0$, it must be positive for all even values of ν.

It remains to show that (35) is satisfied for *odd* ν.

4. First let $1<t<R$. Since ν is odd,

$$q(t)=\frac{(t^2-1)(t^n+1)}{1-2t\cos(\nu\pi/n)+t^2}=\frac{(t^2-1)(t^n+1)}{(t-e^{i\nu\pi/n})(t-e^{-i\nu\pi/n})}$$

is a polynomial of degree n with positive highest coefficient, all of whose roots are on the unit circle $|t|=1$. (For $\nu=n$ the presence of the factor t^2-1 is essential, however not for $1\leqq\nu\leqq n-1$.) Then $q(t)$ is a convex function for $t>1$, and it vanishes at $t=1$. Thus $0<q(t)<q(R)(t-1)/(R-1)$ for $1<t<R$. The expression in curled brackets $\{\ \}$ in (35) is of the form $1-q(t)/q(R)$ and so it must be greater than $(R-t)/(R-1)$, $1<t<R$.

Thus the contribution of the range $1<t<R$ to the integral of (35) will be greater than

$$\frac{R+1}{4}\int_1^R\frac{t^2-1}{t^{1/2}(R-t)^{1/2}(Rt-1)^{3/2}}dt. \tag{37}$$

Now,

$$|(R-t)(Rt-1)|^{1/2}\leqq\tfrac{1}{2}(R^2-1)R^{-1/2},\ |Rt-1|\leqq R^2-1 \text{ for } 1<t<R,$$

so (37) is greater than

$$\tfrac{1}{2}(R+1)^{-1}(R-1)^{-2}\int_1^R(t^2-1)dt=\tfrac{1}{6}(R+2)(R+1)^{-1}>\tfrac{1}{6}.$$

5. Let $-1<t<0$. Referring to (36), we see that in the present case $k(\lambda)$ is increasing (since ν is odd). Then, by what we have just shown, (35) will be a consequence of the following inequality

$$\frac{7}{6}+\int_{-1}^0 p(R;t)\left\{1-\frac{t^2-1}{R^2-1}\,\frac{t^n+1}{R^n+1}\,\frac{(R+1)^2}{(t+1)^2}\right\}dt>0. \tag{38}$$

To prove (38) we divide the integrand into two parts, and estimate them separately. First,

$$-\int_{-1}^{0} p(R;t)dt = \tfrac{1}{4}(R^2-1)\int_0^1 \frac{1-t^2}{t^{1/2}(R+t)^{3/2}(Rt+1)^{3/2}}\,dt$$

$$< \tfrac{1}{4}(R^2-1)R^{-3/2}\int_0^1 t^{-1/2}(Rt+1)^{-3/2}dt$$

$$= \tfrac{1}{2}(R^2-1)R^{-3/2}(R+1)^{-1/2}$$

$$< \tfrac{1}{2}(R^2-1)R^{-2} < \tfrac{1}{2}(R^2-1)/(R^2-R+1).$$

The absolute value of the remaining portion of the integral of (38) can be no larger than

$$\frac{(R+1)^2}{4(R^n+1)}\int_0^1 \frac{(1+t)^2(1+t^n)}{t^{1/2}(R+t)^{3/2}(Rt+1)^{3/2}}\,dt. \tag{39}$$

The integrand will be largest when $R=1$. Then, since $n\geqq 3$, (39) cannot exceed

$$\frac{(R+1)^2}{4(R^3+1)}\int_0^1 \frac{1+t^3}{t^{1/2}(1+t)}\,dt = \frac{13}{30}\,\frac{R+1}{R^2-R+1}.$$

6. Our assertion will now follow from the inequality

$$\tfrac{1}{2}(R^2-1)+\tfrac{13}{30}(R+1) < \tfrac{7}{6}(R^2-R+1), \qquad R>1, \tag{40}$$

which is easily verified. This proves (33).

The non-trivial character of (33) becomes clear from the remark that the inequality

$$\sum_{m=0}^{n}{}' c_{n-m}(R)\cos m\theta \geqq 0$$

is not true for all real values of θ and $R\geqq 1$. Indeed [cf. Part I, (21)]

$$\sum_{m=0}^{n}{}' c_{n-m}(1)\cos m\theta = \sum_{m=0}^{n}{}'\cos m\theta = \tfrac{1}{2}\sin n\theta\cot(\theta/2).$$

This expression changes its sign at $\theta=\nu\pi/n$, $1\leqq\nu\leqq n$.

7. Since $C_{\nu n}>0$, $0\leqq\nu\leqq n$, the equality sign in our main inequalities holds only for the special harmonic functions pointed out in the theorems of Part I. In order to discuss the conditions of equality in (3) we take into account the preliminary remarks of §1. Then we see that for $x=y=0$, $z=R$ the equality holds only if $(2\pi)^{-1}\int_{-\pi}^{\pi}U(1,\theta,\phi)d\phi \equiv \pm\cos n\theta$. Since $|U|\leqq 1$ we conclude (cf. [7, p. 58]) that $U(1,\theta,\phi)\equiv\pm\cos n\theta$.

6. Discussion of $c_n'(1)$ and $c_n''(1)$

1. According to the results of the previous section, we have for $1\leqq\nu\leqq n$ $C_{\nu n}(R)>0$, $R>1$; $C_{\nu n}(1)=0$; so $C_{\nu n}'(1)\geqq 0$—that is,

$$\sum_{m=0}^{n}{}' c'_{n-m}(1)\cos(m\nu\pi/n) \geqq 0, \qquad 1 \leqq \nu \leqq n. \tag{41}$$

This is trivial for $\nu=0$ because $c_n'(1)>0$. The constants $c_n'(1)=\rho_n$ are identical with the constants introduced by Szegö [7, p. 60]. Inequalities (41) furnish the following theorem:

Let $u(x, y, z)$ be a harmonic polynomial of degree n satisfying the condition $|u(x, y, z)| \leqq 1$ in $x^2+y^2+z^2 \leqq 1$. Then on the unit sphere we have

$$\left|\frac{\partial u}{\partial r}\right| \leqq c_n'(1) = \rho_n. \tag{42}$$

If the boundary values of u on the unit sphere are of the form $\pm\cos n\gamma$, γ being the spherical distance of the variable point from a fixed point of the unit sphere, then the equality occurs in (42).

This is slightly less than the main theorem of [7]. It is shown there that (42) is an equality only if $u \equiv \pm\cos n\gamma$.

The representation

$$\rho_n = c_n'(1) = \tfrac{1}{2}\pi n - \tfrac{1}{2} + \frac{(-1)^n}{2}\int_0^1 \frac{(1-t)t^{n-1/2}}{(1+t)^2}\,dt \tag{43}$$

is the analogue of (30). It follows from (30) by differentiation. Another way to obtain (43) is to calculate

$$\left(\frac{\partial}{\partial R}F(R, z)\right)_{R=1} = \left(\frac{\partial}{\partial R}G(R, z)\right)_{R=1} = \sum_{n=1}^{\infty}\rho_n z^n,$$

and to apply to this function considerations similar to those in §§3, 4. This generating function was used in [7].

A direct proof of inequalities (41) is possible based on the representation (43).

2. Now let $\tau_n=\frac{1}{2}c_n''(1)$. We discuss the following inequalities, $n \geqq 2$,

$$\frac{1}{2}C''_{\nu n}(1) = \frac{1}{2}\sum_{m=0}^{n}{}' c''_{n-m}(1)\cos(m\nu\pi/n) = \sum_{m=0}^{n}{}' \tau_{n-m}\cos(m\nu\pi/n) > 0, \tag{44}$$
$$0 \leqq \nu \leqq n.$$

It was pointed out in §1.3 that $c_n''(1)>0$; so inequality (44) is trivial for $\nu=0$. Let $1 \leqq \nu \leqq n$. It is advisable to calculate

$$\left(\frac{\partial^2}{\partial R^2}F(R, z)\right)_{R=1} = \left(\frac{\partial^2}{\partial R^2}G(R, z)\right)_{R=1} + \sum_{n=1}^{\infty} n^2 z^n = 2\sum_{n=2}^{\infty}\tau_n z^n.$$

Introducing the abbreviation $(R-z)(1-Rz)=K$ we obtain

$$\frac{\partial}{\partial R}G(R,z) = (1-z^2)z^{1/2}\left\{RK^{-3/2} - \tfrac{3}{4}(R^2-1)K^{-5/2}(1-2Rz+z^2)\right\} \arctan\left(\frac{z(R-z)}{1-Rz}\right)^{1/2} + (1-z^2)z^{1/2}\tfrac{1}{2}(R^2-1)K^{-3/2}\tfrac{1}{2}z^{1/2}K^{-1/2},$$

so (cf. [7, p. 60, (28)])

$$\left(\frac{\partial}{\partial R}G(R,z)\right)_{R=1} = \sum_{n=1}^{\infty}\rho_n z^n = z^{1/2}(1-z)^{-2}(1+z)\arctan(z^{1/2}).$$

Furthermore

$$\frac{\partial^2}{\partial R^2}G(R,z) = (1-z^2)z^{1/2}\left\{K^{-3/2} - 3RK^{-5/2}(1-2Rz+z^2)\right\}\arctan\left(\frac{z(R-z)}{1-Rz}\right)^{1/2} + (1-z^2)z^{1/2}RK^{-3/2}\tfrac{1}{2}z^{1/2}K^{-1/2} + (1-z^2)z^{1/2}RK^{-3/2}\tfrac{1}{2}z^{1/2}K^{-1/2} + \cdots$$

where the terms which are not written out vanish for $R=1$. This furnishes

$$\left(\frac{\partial^2}{\partial R^2}G(R,z)\right)_{R=1} = -2z^{1/2}(1-z)^{-2}(1+z)\arctan(z^{1/2}) + z(1-z)^{-2}(1+z),$$

or

$$\sum_{n=1}^{\infty}(2\tau_n - n^2)z^n = -2\sum_{n=1}^{\infty}\rho_n z^n + \sum_{n=1}^{\infty}n^2 z^n;$$

so, on account of (43),

$$\tau_n = \tfrac{1}{2}c_n''(1) = -\rho_n + n^2$$

$$= n^2 - \tfrac{1}{2}\pi n + \tfrac{1}{2} + \frac{(-1)^{n+1}}{2}\int_0^1 \frac{(1-t)t^{n-1/2}}{(1+t)^2}\,dt. \tag{45}$$

3. By use of formula (21) of Part I we find

$$\sum_{m=0}^{n}{}'m^2\cos m\theta = -\frac{1}{2}\left(\frac{d}{d\theta}\right)^2(\sin n\theta\cot(\theta/2)).$$

Then for $1 \leqq \nu \leqq n$ we have

$$\sum_{m=0}^{n}{}'(n-m)^2\cos(m\nu\pi/n) = \sum_{m=0}^{n}{}'m^2\cos(\nu\pi - m\nu\pi/n) = (-1)^{\nu}\sum_{m=0}^{n}{}'m^2\cos(m\nu\pi/n)$$

$$= (-1)^{\nu}\left(\frac{n\cos n\theta}{2\sin^2(\theta/2)}\right)_{\theta=\nu\pi/n} = \frac{n}{2\sin^2(\nu\pi/(2n))}.$$

Also

$$\sum_{m=0}^{n}{}' (n-m)\cos(m\nu\pi/n) = \frac{1}{2}\left(\frac{\sin^2(n\theta/2)}{\sin^2(\theta/2)}\right)_{\theta=\nu\pi/n} = \frac{1-(-1)^\nu}{4\sin^2(\nu\pi/(2n))},$$

so from (45) [cf. Part I, (21), (20)] we obtain

$$\sum_{m=0}^{n}{}' \tau_{n-m}\cos m\theta = \frac{n-\frac{1}{4}\pi(1-(-1)^\nu)}{2\sin^2(\theta/2)} - \frac{1}{4}\int_0^1 \frac{(1-t)t^{-1/2}}{(1+t)^2}\,\frac{1-t^2}{1+2t\cos\theta+t^2}\,((-1)^\nu-(-t)^n)\,dt \tag{46}$$

where $\theta=\nu\pi/n$, $1\leqq\nu\leqq n$.

Let ν be *even*. The first term as a function of $\cos\theta$ is increasing; the integral in the second term is decreasing. Therefore the total expression is greater than

$$n/2 - \frac{1}{2}\int_0^1 \frac{t^{-1/2}}{1+t}\,dt = n/2 - \pi/4$$

which is positive for $n\geqq 2$.

Let ν be *odd*. Then, the integral in the second term of (46) is negative and $n-\pi/2>0$, so the statement is clear. Thus, we have proved the following

Let $u(x, y, z)$ be a harmonic polynomial of degree n satisfying the condition $|u(x, y, z)|\leqq 1$ in $x^2+y^2+z^2\leqq 1$. Then on the unit sphere

$$\frac{1}{2}\left|\frac{\partial^2 u}{\partial r^2}\right| \leqq \frac{1}{2}c_n''(1) = \tau_n$$

with the sign "=" if and only if the boundary values of u on the unit sphere are of the form $\pm\cos n\gamma$, γ being the spherical distance of the variable point from a fixed point on the unit sphere.

7. Asymptotic formula of $c_n(R)$ as $n\to\infty$

From (30) and (30′) we obtain for fixed $R>1$ and $n\to\infty$,

$$\begin{aligned} c_n(R) &= O(R^n) + \int_1^R p(R;t)(R^n-t^n)\,dt \\ &= O(R^n) + \frac{(R^2-1)^{1/2}}{4R^{1/2}}\int_1^R \frac{R^n-t^n}{(R-t)^{3/2}}\,dt + O(1)\int_1^R \frac{R^n-t^n}{(R-t)^{1/2}}\,dt \\ &= O(R^n) + \frac{(R^2-1)^{1/2}}{4R^{1/2}}\int_0^R \frac{R^n-t^n}{(R-t)^{3/2}}\,dt \\ &= O(R^n) + \frac{(R^2-1)^{1/2}}{4R}R^n\int_0^1 \frac{1-t^n}{(1-t)^{3/2}}\,dt. \end{aligned}$$

But

$$\int_0^1 \frac{1-t^n}{(1-t)^{3/2}}\,dt = \sum_{m=0}^{n-1}\int_0^1 \frac{t^m}{(1-t)^{1/2}}\,dt = 2\sum_{m=0}^{n-1}\frac{2\cdot 4\cdots 2m}{3\cdot 5\cdots(2m+1)}$$

$$\simeq \sum_{m=1}^{n-1}\pi^{1/2}m^{-1/2} \simeq 2\pi^{1/2}n^{1/2}, \qquad n\to\infty,$$

which furnishes (6).

Formula (6) also follows directly from (3).

Appendix

1. Generalization

A problem similar to that treated in Part II can be formulated for harmonic polynomials in any euclidean space. By using classical results on the representation of such harmonic polynomials in terms of ultraspherical polynomials(16), we are led to the following problem:

Let $P_n^{(\lambda)}(x)$ denote the ultraspherical polynomials, $\lambda>0$, and let R be fixed, $R>1$. Considering all cosine polynomials

$$u(\theta) = \sum_{m=0}^{n} a_m P_m^{(\lambda)}(\cos\theta) \tag{1}$$

of degree $n\geqq 1$, satisfying the condition $|u(\theta)|\leqq 1$ for real θ, what is the maximum of $\left|\sum_{m=0}^{n} a_m P_m^{(\lambda)}(1)R^m\right|$?

1. In case of harmonic polynomials of $p+1$ variables we have to take $\lambda=\frac{1}{2}(p-1)$. However the problem mentioned has sense for arbitrary $\lambda>0$. The two-dimensional case (Part I) corresponds to the limiting value $\lambda\to 0$, the three-dimensional case (Part II) to the value $\lambda=1/2$.

Again the maximum in question is attained for $u(\theta)=\pm\cos n\theta$ if and only if the following condition is satisfied. Let

$$\cos n\theta = \sum_{m=0}^{n} h_{nm}P_m^{(\lambda)}(\cos\theta), \qquad c_n^{(\lambda)}(R) = \sum_{m=0}^{n} h_{nm}P_m^{(\lambda)}(1)R^m; \tag{2}$$

then

$$\sum_{m=0}^{n}{}' c_{n-m}^{(\lambda)}(R)\cos(m\nu\pi/n) \geqq 0, \qquad 0\leqq\nu\leqq n. \tag{3}$$

2. The proof of this inequality for general λ seems to be rather difficult.

(16) See, for instance, E. Heine, *Handbuch der Kugelfunctionen*, vol. 1 (2d edition, 1878); see in particular pp. 454 and 461. Regarding the notation of ultraspherical polynomials we follow Szegö's book quoted in Part II, §2; see in particular p. 80.

In the limiting case $\lambda \to 0$ cf. formula (24) of Part I. In the case $\lambda = 1/2$ we have inequality (33) of Part II.

Another simple case is $\lambda = 1$. Since

$$\cos n\theta = \frac{1}{2}\frac{\sin (n+1)\theta}{\sin \theta} - \frac{1}{2}\frac{\sin (n-1)\theta}{\sin \theta}, \qquad n \geqq 1, \tag{4}$$

we have

$$c_n^{(1)}(R) = \tfrac{1}{2}(n+1)R^n - \tfrac{1}{2}(n-1)R^{n-2}, \qquad n \geqq 1; \qquad c_0^{(1)}(R) = 1, \tag{5}$$

so that

$$\begin{aligned}
&\sum_{m=0}^{n}{}' c_{n-m}^{(1)}(R)\cos(m\nu\pi/n) \\
&\quad = \frac{1}{2}\frac{d}{dR}\left\{(R - R^{-1})\sum_{m=0}^{n}{}' R^{n-m}\cos(m\nu\pi/n)\right\} \\
&\quad\quad - \tfrac{1}{4}(-1)^\nu \frac{d}{dR}(R - R^{-1}) + \tfrac{1}{2}(-1)^\nu \qquad (6) \\
&\quad = \tfrac{1}{2}(1 + R^{-2})\sum_{m=0}^{n}{}' R^{n-m}\cos(m\nu\pi/n) \\
&\quad\quad + \tfrac{1}{2}(1 - R^{-2})\sum_{m=0}^{n}{}' (n-m)R^{n-m}\cos(m\nu\pi/n) + \tfrac{1}{4}(-1)^\nu(1 - R^{-2}).
\end{aligned}$$

According to well known properties of Fejér's kernel, the second sum of the last expression is positive. The first sum is [cf. Part I, (20)]

$$\tfrac{1}{2}\frac{(R^2-1)(R^n - (-1)^\nu)}{1 - 2R\cos(\nu\pi/n) + R^2}.$$

This furnishes the positivity of (6) when ν is even. Let ν be odd; then we have only to show that

$$(1 + R^{-2})\frac{(R^2-1)(R^n+1)}{1 - 2R\cos(\nu\pi/n) + R^2} > 1 - R^{-2},$$

which is clear.

Thus we established the following theorem:

Let $n \geqq 1$ and let

$$u(\theta) = \sum_{m=0}^{n}\beta_m \frac{\sin (m+1)\theta}{\sin \theta} \tag{7}$$

be an arbitrary cosine polynomial of degree n satisfying the condition $|u(\theta)| \leqq 1$, *θ real. Then for* $R>1$

$$\left|\sum_{m=0}^{n}(m+1)\beta_m R^m\right| \leqq \tfrac{1}{2}(n+1)R^n - \tfrac{1}{2}(n-1)R^{n-2} \tag{8}$$

with the sign "=" if and only if $u(\theta) = \pm\cos n\theta$.

2. A PROBLEM ON ELLIPSES

1. Some time ago W. E. Sewell[11] dealt with the following extremum problem:

Let E be an ellipse in the complex z-plane. If $f(z)$ *is a polynomial of degree n satisfying the condition* $|f(z)| \geqq 1, z\in E$, *what is the maximum of* $\max |f'(z)|$, $z\in E$?

Sewell obtains an upper bound for this maximum, which is, however, not the precise one. His bound is of order n as $n\to\infty$, which is the precise order[12].

2. For sake of convenience we consider an ellipse E_R with foci at -1 and $+1$ and with semi-axes $\frac{1}{2}(R\pm R^{-1})$, $R>1$. Let $\alpha=\frac{1}{2}(R+R^{-1})$. We prove the following result of *negative nature*:

Let $n\geqq 5$. *There exist polynomials* $f(z)$ *of degree n satisfying the condition* $|f(z)| \leqq 1, z\in E_R$, *such that*

$$|f'(\alpha)| > \frac{2n}{R-R^{-1}}\,\frac{R^n-R^{-n}}{R^n+R^{-n}} = \frac{T_n'(\alpha)}{T_n(\alpha)}. \tag{9}$$

Here $T_n(z)$ is Tchebychef's polynomial; obviously $|T_n(z)/T_n(\alpha)| \leqq 1$ when $z\in E_R$. This shows that for $n\geqq 5$ the solution of the problem mentioned is *not* furnished by Tchebychef's polynomial. This is remarkable since Tchebychef's polynomial does furnish the maximum in question for all n in the limiting case $R\to 1$, that is, in the case of the segment $(-1, +1)$.

3. Let $2z=Re^{i\phi}+R^{-1}e^{-i\phi}$. If ϕ runs from 0 to 2π, z describes E_R; $\phi=0$ furnishes $z=\alpha$. We have with certain complex coefficients c_m

$$\begin{aligned} f(z) &= \sum_{m=0}^{n} c_m(R^m e^{im\phi} + R^{-m}e^{-im\phi}) \\ &= \sum_{m=0}^{n} c_m(R^m + R^{-m})\cos m\phi + i\sum_{m=0}^{n} c_m(R^m - R^{-m})\sin m\phi, \qquad (10) \\ f'(\alpha) &= \sum_{m=0}^{n} 2mc_m\,\frac{R^m-R^{-m}}{R-R^{-1}}. \end{aligned}$$

[11] W. E. Sewell, *On the polynomial derivative constant for an ellipse*, American Mathematical Monthly, vol. 44 (1937), pp. 577–578.

[12] G. Szegö, *Über einen Satz von A. Markoff*, Mathematische Zeitschrift, vol. 23 (1925), pp. 45–61; see in particular p. 53.

Substituting in the interpolation formula of Part I, §2:

$$\lambda_m = \frac{2m}{R - R^{-1}} \frac{R^m - R^{-m}}{R^m + R^{-m}}, \qquad \mu_m = 0, \qquad 0 \leqq m \leqq n, \tag{11}$$

we obtain

$$f'(\alpha) = \frac{1}{2n} \sum^{2n-1} (-1)^\nu G_\nu f(z_\nu), \qquad z_\nu = \tfrac{1}{2}(Re^{i\nu\pi/n} + R^{-1}e^{-i\nu\pi/n}),$$
$$G_\nu = \frac{4}{R - R^{-1}} \sum^{n}{}' (n - m) \frac{R^{n-m} - R^{-(n-m)}}{R^{n-m} + R^{-(n-m)}} \cos(m\nu\pi/n). \tag{12}$$

4. Our next purpose is to show that for $n \geqq 5$, for some $\nu = \nu(n, R)$ and for sufficiently large values of R, the inequality $G_\nu < 0$ holds. Indeed as $R \to \infty$, $1 \leqq \nu \leqq n$,

$$\tfrac{1}{4}(R - R^{-1})G_\nu = \sum_{m=0}^{n}{}' (n - m) \cos(m\nu\pi/n) - 2R^{-2} \cos((n-1)\nu\pi/n) + O(R^{-4})$$
$$= \frac{1}{2} \frac{1 - (-1)^\nu}{1 - \cos(\nu\pi/n)} - 2R^{-2} \cdot (-1)^\nu \cos(\nu\pi/n) + O(R^{-4}). \tag{13}$$

Let ν be even and $0 < \nu < n/2$; then the last expression becomes negative if R is sufficiently large. The condition regarding ν implies $n/2 > 2$, $n \geqq 5$, and indeed for $n \geqq 5$ values $\nu = \nu(n)$ of the kind mentioned always exist.

5. Finally let $n \geqq 5$, $G_{\nu(n)} < 0$ (R sufficiently large) and

$$f(z) = T_n(z)/T_n(\alpha) - \eta h(z) \tag{14}$$

where $\eta > 0$ and $h(z)$ is a polynomial of degree n satisfying the following conditions:

(a) that $h(z)$ has real coefficients;

(b) that $(-1)^\nu \Re h(z_\nu) > 0$, $0 \leqq \nu \leqq 2n-1$;

(c) that

$$h'(\alpha) = \frac{1}{2n} \sum_{\nu=0}^{2n-1} (-1)^\nu G_\nu h(z_\nu)$$
$$= \frac{1}{2n} \Re\left\{G_0 h(z_0) + 2\sum_{\nu=1}^{n-1} (-1)^\nu G_\nu h(z_\nu) + (-1)^n G_n h(z_n)\right\} < 0. \tag{15}$$

In (15) we have taken into account that $G_{n+\mu} = G_{n-\mu}$, $z_{n+\mu} = \bar{z}_{n-\mu}$, which follows from (12).

Such a polynomial $h(z)$ can easily be constructed. We write if $z \in E_R$

$$h(z) = \sum_{m=0}^{n} h_m(R^m e^{im\phi} + R^{-m}e^{-im\phi}); \tag{16}$$

the coefficients h_m being real,

$$\Re h(z_\nu) = \sum_{m=0}^{n} h_m(R^m + R^{-m}) \cos(m\nu\pi/n). \tag{17}$$

Now we determine the real coefficients h_m such that $\Re h(z_\nu)$ assumes preassigned values for $0 \leqq \nu \leqq n$, in particular such that (b) is satisfied for $0 \leqq \nu \leqq n$; then the same condition (b) is satisfied for $0 \leqq \nu \leqq 2n-1$ since $z_{n+\mu} = \bar{z}_{n-\mu}$. Finally (c) follows if the further condition is imposed that $(-1)^\nu \Re h(z_\nu)$ should be so large for $\nu = \nu(n)$ as to furnish the dominant term in $h'(\alpha)$.

Since $\max |T_n(z)|$, $z \in E_R$, is attained only for $z = z_\nu$, $0 \leqq \nu \leqq 2n-1$, we obtain by a familiar argument, $|f(z)| \leqq 1$, $z \in E_R$, provided η is sufficiently small. From

$$f'(\alpha) = T_n'(\alpha)/T_n(\alpha) - \eta h'(\alpha)$$

we conclude that $|f'(\alpha)| > T_n'(\alpha)/T_n(\alpha)$ provided η is sufficiently small.

This establishes the result stated.

Bibliography

1. S. Bernstein, *Sur un théorème de M. Szegö*, Prace Matematyczno-Fizyczne, vol. 44 (1937), pp. 9–14.

2. M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 23 (1915), pp. 354–368.

3. G. Szegö, *Über trigonometrische und harmonische Polynome*, Mathematische Annalen, vol. 79 (1919), pp. 323–339.

4. ———, *Über einen Satz des Herrn S. Bernstein*, Schriften der Königsberger Gelehrten Gesellschaft, 1928, pp. 59–70.

5. ———, *Über trigonometrische Interpolation*, Schriften der Königsberger Gelehrten Gesellschaft, 1928, pp. 71–80.

6. ———, *Solution to Problem 3705 (proposed by Raphael Robinson)*, American Mathematical Monthly, vol. 43 (1936), pp. 246–259.

7. ———, *On the gradient of solid harmonic polynomials*, these Transactions, vol. 47 (1940), pp. 51–65.

Stanford University,
Stanford University, Calif.

FOURIER INTEGRALS AND METRIC GEOMETRY

BY

J. VON NEUMANN AND I. J. SCHOENBERG

INTRODUCTION

1. Let S be a metric space, the distance $d(f, g)$ between two elements of S satisfying the usual postulates. Let f_t $(-\infty < t < \infty)$ be a function whose values lie in S and which is metrically continuous, that is, $d(f_{t+h}, f_t) \to 0$ if $h \to 0$. If this function has the property that

$$d(f_t, f_s) = F(t - s) \tag{0.1}$$

is a function of the difference $t-s$ only, we call the curve Γ in S defined by f_t $(-\infty < t < \infty)$ a *screw line* of S and $F(t) = d(f_t, f_0)$ a *screw function* of S. The reason for this terminology is as follows. If τ is a real parameter, the two curves $\Gamma_0: f_t$ $(-\infty < t < \infty)$ and $\Gamma_\tau: f_{t+\tau}$ $(-\infty < t < \infty)$ which are identical as point sets, are isometrically mapped on each other by the correspondence $f_t \leftrightarrow f_{t+\tau}$, for

$$d(f_t, f_s) = F(t - s) = d(f_{t+\tau}, f_{s+\tau}),$$

in view of (0.1). These congruent mappings of Γ into itself form a one-parameter group.

The following properties of a screw function $F(t)$ of S are obvious: $F(t)$ is a continuous non-negative even function and $F(0) = 0$; if $F(t)$ is a screw function, then all functions $F(kt)$ (k real) are screw functions.

A different point of view which puts the emphasis on the screw function $F(t)$ rather than on the screw line Γ is as follows. Consider the real axis $-\infty < t < \infty$ as a euclidean space E_1 and change its metric from $|t-s|$ to $F(t-s)$. We thus get a new space which, following Blumenthal, we shall call the metric transform of E_1 by $F(t)$ and denote by $F(E_1)$. For what functions $F(t)$ may this metric transform $F(E_1)$ be isometrically imbedded in S? Clearly $F(E_1)$ enjoys this property if and only if $F(t)$ is a screw function of S. For if the mapping of the point t, of $F(E_1)$, into the point f_t, of S, performs the imbedding of $F(E_1)$ into S, then (0.1) expresses the isometricity of this imbedding.

Presented to the Society, September 1, 1936, and December 29, 1938; received by the editors September 23, 1940. The first communication to the Society [Bulletin of the American Mathematical Society, abstract 42-9-353] covered the contents of Part I and Part II of the present paper; the contents of Part III were the subject of the second communication of 1938 [abstract 47-1-47]. Thus the contents of Parts I and II precede in time the articles [4] and [5], listed in the bibliography at the end of this paper, which were partly suggested by this earlier work frequently referred to in these articles. Part III, however, carries on the work presented in [4] and [5] and is essentially based on some of that work.

2. As examples we mention

$$F_3(t) = (t^2 + \sin^2 t)^{1/2},$$

which is a screw function of the euclidean space E_3, in view of the identity

$$F_3^2(t-s) = (t-s)^2 + \sin^2(t-s)$$
$$= (t-s)^2 + \tfrac{1}{2}(\cos 2t - \cos 2s)^2 + \tfrac{1}{2}(\sin 2t - \sin 2s)^2.$$

Similarly

$$F_{2N}(t) = \left(\sum_{\nu=1}^{N} A_\nu \sin^2 u_\nu t\right)^{1/2}, \qquad A_\nu > 0,\ 0 < u_1 < u_2 < \cdots < u_N, \tag{0.2}$$

is a screw function of E_{2N}, as seen from

$$F_{2N}^2(t-s) = \sum_{\nu=1}^{N} \{\tfrac{1}{2}A_\nu(\cos 2u_\nu t - \cos 2u_\nu s)^2 + \tfrac{1}{2}A_\nu(\sin 2u_\nu t - \sin 2u_\nu s)^2\}. \tag{0.2'}$$

Finally

$$F_{2N-1}(t) = \left(Ct^2 + \sum_{\nu=1}^{N-1} A_\nu \sin^2 u_\nu t\right)^{1/2}, \tag{0.3}$$
$$C > 0,\ A_\nu > 0,\ 0 < u_1 < \cdots < u_{N-1},$$

is a screw function of E_{2N-1}, as shown by

$$F_{2N-1}^2(t-s) = C(t-s)^2 + \sum_{\nu=1}^{N-1} \{\tfrac{1}{2}A_\nu(\cos 2u_\nu t - \cos 2u_\nu s)^2 + \tfrac{1}{2}A_\nu(\sin 2u_\nu t - \sin 2u_\nu s)^2\}. \tag{0.3'}$$

We shall see that (0.2) and (0.3) are the most general screw functions of E_{2N} and E_{2N-1}, respectively, which are not also screw functions of a euclidean space of lower dimensions.

The starting point of the present investigation was W. A. Wilson's recent remark that if

$$F(t) = |t|^{1/2} \tag{0.4}$$

then the metric transform $F(E_1)$ is imbeddable in the Hilbert space $\mathfrak{H}$, hence

$$F^2(t-s) = |t-s| = \|f_t - f_s\|^2,$$

for a suitable function f_t $(-\infty < t < \infty)$ with values in $\mathfrak{H}$ [8, p. 64]. According to our present point of view (0.4) is a screw function of $\mathfrak{H}$. We shall see that this $F(t)$ is not a screw function of any euclidean space.

3. The principal purpose of this paper is to determine all screw functions of Hilbert space. It consists of three parts. In Part I we state our fundamental

result (Theorem 1) and show by means of elementary results of Menger and Schoenberg that all functions $F(t)$ there described are screw functions of $\mathfrak{H}$. Furthermore, those screw functions of $\mathfrak{H}$ are characterized which correspond to screw lines Γ with one of the following properties: Γ is euclidean, bounded, rectifiable or closed.

The converse statement to the effect that Theorem 1 yields *all* screw functions of $\mathfrak{H}$ is established in two essentially different ways in Part II and Part III respectively. In Part II this is proved by a direct investigation of the group of isometric mappings of $\mathfrak{H}$ into itself which is induced by the group of isometric mappings of a screw line into itself. Free use is made of the theory of Hermitian operators in Hilbert space. In Part III $\mathfrak{H}$ intervenes only by its metric in accordance with the ideas of Menger on the metric characterization of metric spaces. The method used is an elaboration of the metrical approach of Part I which is made more effective by an appeal to the theory of positive definite functions, i.e., the characteristic functions of the theory of probabilities. This connection was pointed out in two recent papers by one of us [4, 5]. In [5, p. 837], it was shown that the question as to when the metric transform $F(E_m)$ is imbeddable in $\mathfrak{H}$ depended on certain limit (closure) theorems. Here we establish these theorems in a form similar to P. Lévy's limit theorem concerning characteristic functions.

In conclusion we want to say that the operator method of Part II, dealing with the imbedding of $F(E_1)$ in $\mathfrak{H}$, may also be extended to cover the general case of the imbedding of $F(E_m)$ in $\mathfrak{H}$. Although this extension has been fully worked out, for reasons of conciseness we treat in Part II only the case of screw lines ($m=1$).

Part I. The fundamental theorem on screw functions in Hilbert space and elementary consequences

1.1. Let $\mathfrak{H}$ be a real Hilbert space. For every $t>-\infty$, and $<+\infty$, let a point f_t of $\mathfrak{H}$ be given, such that

(i) f_t is a metrically continuous function of t,

(ii) the distance of f_t and f_s depends on $t-s$ only.

According to the general definition of our Introduction, the curve $\Gamma: f_t$ is a *screw line* of $\mathfrak{H}$. Condition (ii) means

$$\|f_t - f_s\| = F(t-s). \tag{1.1}$$

Then (i) merely states that

$$F(t) \to 0 \quad \text{if} \quad t \to 0, \tag{1.2}$$

and conversely, (1.2) implies the continuity of f_t and therefore of $F(t)$ everywhere. $F(t)$ is called a *screw function* of $\mathfrak{H}$.

Let $f_{t_0}, f_{t_1}, \cdots, f_{t_n}$ ($t_0=0$) be $n+1$ points of Γ. By (1.1) their mutual distances are $\|f_{t_i}-f_{t_k}\| = F(t_i-t_k)$. As these points may be thought of as lying

in a n-dimensional linear subspace of $\mathfrak{H}$, i.e., a E_n, we have by a known theorem [3, Theorem 1]

$$\sum_{i,k=1}^{n} \{F^2(t_i) + F^2(t_k) - F^2(t_i - t_k)\}\rho_i\rho_k \geqq 0, \tag{1.3}$$

for arbitrary real ρ_i. Conversely, let $F(t)$ be a continuous even non-negative function, with $F(0)=0$, enjoying the property (1.3) for arbitrary real t_i, ρ_i $(n=2, 3, \cdots)$. Geometrically this means: The points $t_0=0, t_1, \cdots, t_n$, of the space $F(E_1)$, which is obtained from E_1 by changing its metric from $|t-s|$ to $F(t-s)$, may be imbedded in E_n. By a theorem of Menger [2] this is sufficient to insure the possibility of imbedding $F(E_1)$ in $\mathfrak{H}$. We can therefore state

LEMMA 1. *A continuous even non-negative function $F(t)$ vanishing at the origin is a screw function of $\mathfrak{H}$ if and only if it satisfies the inequality* (1.3) *for arbitrary real t_i, ρ_i and for $n=2, 3, 4, \cdots$.*

The interest of this analytical characterization of screw function lies in its obvious consequence that the squares $F^2(t)$ of screw functions form a convex class of functions: If $F_1^2(t)$ and $F_2^2(t)$ are squares of screw functions then also $F_1^2(t)+F_2^2(t)$ is the square of a screw function. In view of the euclidean screw functions (0.2), (0.3), exhibited in the Introduction, this convexity property suggests the following explicit expression of the screw functions of $\mathfrak{H}$.

THEOREM 1. (Fundamental theorem.) *The class of screw functions $F(t)$ of Hilbert space is identical with the class of functions whose squares are of the form*

$$F^2(t) = \int_0^\infty \frac{\sin^2 tu}{u^2}\, d\gamma(u), \tag{1.4}$$

where $\gamma(u)$ is non-decreasing for $u \geqq 0$ and such that

$$\int_1^\infty u^{-2} d\gamma(u) \text{ exists.} \tag{1.5}$$

A proof that all functions $F(t)$ furnished by (1.4) are screw functions is exceedingly simple. For it suffices to show that $F^2(t)$ satisfies the inequality (1.3) (Lemma 1). Indeed, in view of the identity

$$\sin^2 t_i u + \sin^2 t_k u - \sin^2 (t_i - t_k)u = 2 \sin^2 t_i u \sin^2 t_k u + \tfrac{1}{2} \sin 2t_i u \sin 2t_k u,$$

we have

$$\begin{aligned} &\sum_{i,k=1}^{n} \{F^2(t_i) + F^2(t_k) - F^2(t_i - t_k)\}\rho_i\rho_k \\ &= \int_0^\infty \left\{2\left(\sum_1^n \rho_i \sin^2 t_i u\right)^2 + \frac{1}{2}\left(\sum_1^n \rho_i \sin 2t_i u\right)^2\right\} u^{-2} d\gamma(u) \geqq 0. \end{aligned} \tag{1.6}$$

The converse statement that the square of a screw function is necessarily of the form (1.4) is proved in two different ways in Part II and Part III of this paper.

1.2. Let us now see which of the screw functions of $\mathfrak{H}$ are euclidean screw functions, that is, belong to some euclidean space. It is convenient for this purpose to write (1.4) in the form

$$F^2(t) = Ct^2 + \int_{+0}^{\infty} \frac{\sin^2 tu}{u^2}\, d\gamma(u), \qquad C = \gamma(+0) - \gamma(0). \tag{1.7}$$

THEOREM 2. *A screw function $F(t)$ defined by (1.4) is euclidean if and only if $\gamma(u)$ has a finite number N of points of increase. $F(t)$ belongs to E_m and to no lower space $E_{m'}$ $(m' < m)$ if and only if*

$$C = 0, \qquad N = m/2 \qquad \textit{for } m \textit{ even}, \tag{1.8}$$

$$C > 0, \qquad N = (m+1)/2 \qquad \textit{for } m \textit{ odd}. \tag{1.9}$$

Indeed, let N be finite. If $C=0$, then $F(t)$ is of the form (0.2), and $u_1, \cdots, u_N$ are precisely the points of increase of $\gamma(u)$. We know by (0.2′) that $F(t)$ is a screw function of E_{2N}. Moreover its screw line

$$(\tfrac{1}{2}A_\nu^{1/2} \cos 2u_\nu t,\ \tfrac{1}{2}A_\nu^{1/2} \sin 2u_\nu t), \qquad \nu = 1, 2, \cdots, N; \ -\infty < t < \infty,$$

lies in E_{2N} but in no linear subspace of E_{2N}. Similarly, if $C>0$, it is shown by (0.3), (0.3′), that $F(t)$ belongs to E_m, $m=2N-1$.

Let now $\gamma(u)$ have infinitely many points of increase. We want to show that $F(t)$ is not euclidean, that is, the metric transform $F(E_1)$ cannot be imbedded in a euclidean space. For if n is an arbitrary positive integer, we may find n points of increase $u=u_\nu$ $(\nu=1, \cdots, n)$ of $\gamma(u)$, such that $0<u_1<u_2< \cdots <u_n$. We shall now locate in $F(E_1)$ $n+1$ points $t=0$, $t_1, \cdots, t_n$, which cannot be imbedded in E_{n-1}. Such points will enjoy this property if the quadratic form (1.6) is positive definite [3, Theorem 1]. The linear independence of the functions $\sin 2tu_1, \cdots, \sin 2tu_n$, implies the existence of values t_i, such that

$$\det \|\sin (2t_i u_k)\|_{1,n} \neq 0. \tag{1.10}$$

The form (1.6) is now positive definite; for otherwise there would exist values ρ_i, with

$$\sum \rho_i^2 > 0, \tag{1.11}$$

which make the form (1.6) vanish. But the vanishing of the integral of (1.6) implies the vanishing of its integrand at all points of increase of $\gamma(u)$, hence in particular

$$\sum_{i=1}^{n} \rho_i \sin(2t_i u_k) = 0, \qquad k = 1, \cdots, n. \tag{1.12}$$

The incompatibility of the three relations (1.10), (1.11), (1.12) completes this indirect proof; for as n may be taken arbitrarily large, $F(t)$ is not euclidean.

Wilson's screw function (0.4) is of the form (1.4) since

$$F^2(t) = |t| = \frac{2}{\pi}\int_0^\infty \frac{\sin^2 tu}{u^2}\,du. \tag{1.13}$$

Since $\gamma(u) = 2u/\pi$ has infinitely many points of increase, it is not a euclidean screw function. A somewhat more general example is

$$F(t) = |t|^\kappa, \qquad 0 < \kappa < 1. \tag{1.14}$$

Indeed

$$F^2(t) = |t|^{2\kappa} = \int_0^\infty \sin^2 tu \cdot u^{-1-2\kappa}du \Big/ \int_0^\infty \sin^2 u \cdot u^{-1-2\kappa}du, \qquad 0 < \kappa < 1,$$

is of the form (1.4).

1.3. Let us find conditions which insure the boundedness of a screw function $F(t)$. Later we shall see that this occurs when the corresponding screw line lies on a sphere of $\mathfrak{H}$. For the present we prove

THEOREM 3. *The screw function $F(t)$ given by*

$$F^2(t) = Ct^2 + \int_{+0}^\infty \frac{\sin^2 tu}{u^2}\,d\gamma(u), \qquad C \geqq 0, \tag{1.15}$$

is bounded if and only if

$$C = 0 \quad \textit{and} \quad \int_{+0}^\infty u^{-2}d\gamma(u) \quad \textit{exists.} \tag{1.16}$$

More precisely, if $C=0$, we have

$$\frac{1}{2}\int_{+0}^\infty \frac{d\gamma(u)}{u^2} \leqq \limsup_{t\to\infty} F^2(t) \leqq \int_{+0}^\infty \frac{d\gamma(u)}{u^2}, \tag{1.17}$$

where all three members may also be infinite.

As $C=0$ is obviously necessary in order that $F^2(t)$ be bounded, it suffices to prove the inequalities (1.17). Let

$$F^2_{\epsilon,a}(t) = \int_\epsilon^a \frac{\sin^2 tu}{u^2}\,d\gamma(u), \qquad 0 < \epsilon < a < +\infty.$$

We have

$$\frac{1}{T}\int_0^T F_{\epsilon,a}^2(t)dt = \int_\epsilon^a \left(\frac{1}{T}\int_0^T \sin^2 tu\, dt\right) u^{-2}d\gamma(u)$$

$$= \int_\epsilon^a \left(\frac{1}{2} - \frac{\sin 2Tu}{4Tu}\right) u^{-2}d\gamma(u),$$

hence

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T F_{\epsilon,a}^2(t)dt = \frac{1}{2}\int_\epsilon^a u^{-2}d\gamma(u).$$

Whence, since $F^2(t) \geqq F_{\epsilon,a}^2(t)$, we derive

$$\limsup_{t\to\infty} F^2(t) \geqq \limsup_{t\to\infty} F_{\epsilon,a}^2(t) \geqq \frac{1}{2}\int_\epsilon^a u^{-2}d\gamma(u).$$

By comparing the extreme terms only and allowing $\epsilon\to 0$ and $a\to\infty$, we get the first inequality (1.17). The second inequality (1.17) follows from

$$F^2(t) = \int_{+0}^\infty \frac{\sin^2 tu}{u^2}\, d\gamma(u) \leqq \int_{+0}^\infty \frac{1}{u^2}\, d\gamma(u).$$

The euclidean screw functions (0.2) are always bounded; (0.3) are never bounded.

1.4. We shall now investigate when a screw line Γ of a screw function $F(t)$ is rectifiable. With this purpose in view we show first that

$$\lim_{t\to 0}\left(\frac{F(t)}{t}\right)^2 = \int_0^\infty d\gamma(u), \tag{1.18}$$

where both sides may also be infinite.

Indeed, to each $\epsilon>0$ there corresponds a $\delta>0$, such that $\sin x^2/x^2 > 1-\epsilon$ if $x\geqq 0$ and $\leqq\delta$, hence $\sin^2 tu/(tu)^2 > 1-\epsilon$ if $0\leqq u\leqq \delta t^{-1}$, $t>0$. But then

$$\left(\frac{F(t)}{t}\right)^2 = \int_0^\infty \frac{\sin^2 tu}{t^2u^2}\, d\gamma(u) \geqq \int_0^{\delta t^{-1}} \frac{\sin^2 tu}{t^2u^2}\, d\gamma(u) \geqq (1-\epsilon)\int_0^{\delta t^{-1}} d\gamma(u),$$

and therefore

$$\liminf_{t\to 0}\left(\frac{F(t)}{t}\right)^2 \geqq (1-\epsilon)\int_0^\infty d\gamma(u).$$

On allowing $\epsilon\to 0$, we have

$$\liminf_{t\to 0}\left(\frac{F(t)}{t}\right)^2 \geqq \int_0^\infty d\gamma(u).$$

This, and the obvious inequality

$$\left(\frac{F(t)}{t}\right)^2 = \int_0^\infty \frac{\sin^2 tu}{t^2u^2}\, d\gamma(u) \leqq \int_0^\infty d\gamma(u)$$

imply (1.18).

Let now $t_0=0<t_1<\cdots<t_{n-1}<t_n=t$ be a subdivision of the fixed interval $(0, t)$ and $\delta=\max(t_i-t_{i-1})$. (1.1) and (1.18) now imply

$$\lim_{\delta\to 0} \sum_{i=1}^{n} \|f_{t_i} - f_{t_{i-1}}\| = \lim_{\delta\to 0} \sum_{1}^{n} F(t_i - t_{i-1})$$

$$= \lim_{\delta\to 0} \sum_{1}^{n} \frac{F(t_i - t_{i-1})}{t_i - t_{i-1}} \cdot (t_i - t_{i-1}) = t \cdot \left(\int_0^\infty d\gamma(u)\right)^{1/2}.$$

This proves the following theorem.

THEOREM 4. *A screw line f_t $(-\infty<t<\infty)$ corresponding to the screw function $F(t)$ defined by* (1.4) *is rectifiable if and only if $\gamma(u)$ is bounded. The length s of the arc $(0, t)$ of the screw line is then connected with the parameter t by the relation*

$$s = t(\gamma(\infty) - \gamma(0))^{1/2} = t \cdot \lim_{\tau\to 0} \frac{F(\tau)}{\tau}. \tag{1.19}$$

In particular, if $\gamma(\infty)-\gamma(0)=1$, t is identical with the length of arc along the screw line of $F(t)$.

The euclidean screw lines corresponding to (0.2), (0.3) are always rectifiable, the relations between s and t being respectively

$$s = t\left(\sum_1^N A_\nu u_\nu^2\right)^{1/2}, \qquad s = t\left(C + \sum_1^{N-1} A_\nu u_\nu^2\right)^{1/2}.$$

The screw lines of (1.14) are non-rectifiable, since $F(t)/t\to\infty$ as $t\to 0$.

1.5. When is a screw line Γ, corresponding to the screw function $F(t)$, a closed curve? Let f_{t_0} be a double point of Γ, that is, $f_{t_0}=f_{t_0+\tau}$ for some $\tau>0$. Then $0=\|f_{t_0+\tau}-f_{t_0}\|=F(\tau)$, hence $\|f_{t+\tau}-f_t\|=F(\tau)=0$ or $f_{t+\tau}=f_t$ for all real t. This means that f_t and therefore also $F(t)=\|f_t-f_0\|$ has the period τ and Γ is a closed curve. Conversely, all this is implied if $F(t)$ is periodic. Let this be the case and let τ be the least positive period of $F(t)$, which exists if $F(t)\not\equiv 0$. From (1.7) we get

$$F^2(\tau) = C\tau^2 + \int_{+0}^\infty \frac{\sin^2 \tau u}{u^2}\, d\gamma(u) = 0.$$

This implies that $C=0$ and that $\gamma(u)$ is constant in all the intervals

$(k-1)\pi/\tau < u < k\pi/\tau$ $(k=1, 2, 3, \cdots)$ where $\sin^2 \tau u > 0$. But then $F^2(t)$ reduces to the series

$$F^2(t) = \sum_{k=1}^{\infty} c_k \sin^2\left(\frac{k\pi}{\tau} t\right), \qquad \tau > 0,\ c_k \geqq 0,\ \sum_1^{\infty} c_k \text{ convergent.} \tag{1.20}$$

THEOREM 5. *A screw line of $\mathfrak{H}$ is a closed curve if and only if its screw function is periodic and its square is of the form* (1.20).

The screw lines of (1.20) are rectifiable if and only if $\sum_1^{\infty} k^2 c_k$ converges.

A euclidean screw line is closed only if it is even-dimensional, that is, only if the form (0.2) and its frequencies u_ν $(\nu=1, \cdots, N)$ have rational ratios(1).

The Fourier series developments of the Bernoulli polynomials $B_2(t)$, $B_4(t)$ furnish the following simple examples of periodic screw functions of period one:

$$F_1^2(t) = t(1-t) = \frac{1}{6} - B_2(t) = \sum_1^{\infty} \frac{1-\cos 2\pi nt}{n^2\pi^2} = \frac{2}{\pi^2} \sum_1^{\infty} \frac{\sin^2 \pi nt}{n^2}, \tag{1.21}$$

$$0 \leqq t \leqq 1,$$

$$F_2^2(t) = t^2(1-t)^2 = \frac{1}{30} + B_4(t) = 3\sum_1^{\infty} \frac{1-\cos 2\pi nt}{n^4\pi^4} = \frac{6}{\pi^4} \sum_1^{\infty} \frac{\sin^2 \pi nt}{n^4}, \tag{1.21'}$$

$$0 \leqq t \leqq 1.$$

The screw lines of $F_1(t)$ are non-rectifiable; those of $F_2(t)$ are rectifiable with t as length of arc, since $\lim_{t\to 0} F_2(t)/t = 1$.

1.6. We conclude this survey of these most elementary and characteristic features of screw lines by proving the following theorem.

THEOREM 6. *A screw line Γ of $\mathfrak{H}$ is bounded if and only if it can be placed on a sphere of $\mathfrak{H}$.*

The sufficiency of the condition is clear. Conversely, let the screw line and therefore also its screw function $F(t)$ be bounded. By Theorem 3 we have

$$F^2(t) = \int_{+0}^{\infty} \frac{\sin^2 tu}{u^2}\, d\gamma(u).$$

Setting

$$\alpha(0) = 0, \qquad \alpha(u) = \int_{+0}^{u} u^{-2} d\gamma(u), \qquad u > 0,$$

(1) The projections of such screw lines on any of the coordinate planes $x_\mu O x_\lambda$ are closed Lissajou curves. Conversely, any Lissajou curve whether open or closed may be regarded as a plane projection of a screw line of E_4.

we get

$$F^2(t) = \int_0^\infty \sin^2 tu \, d\alpha(u), \tag{1.22}$$

where $\alpha(u)$ is non-decreasing and bounded. Define $r>0$ by

$$4r^2 = \alpha(\infty). \tag{1.23}$$

We shall prove that Γ can be placed on a sphere of $\mathfrak{H}$ of radius r. Consider the metric transform $F(E_1)$, i.e., the real axis $-\infty<t<\infty$ with the distance $d(t, t') = F(t-t')$; add to this space a new point A and complete the definition of distance throughout $F(E_1)+A$ by setting

$$d(A, t) = r. \tag{1.24}$$

Not only $F(E_1)$ can be inbedded in $\mathfrak{H}$, as shown by (1.6), but also $F(E_1)+A$ may so be imbedded. It suffices to show that the points $A, t_1, \cdots, t_n$ may be placed in E_n. Indeed, we have in view of (1.22), (1.23), (1.24)

$$\begin{aligned}
&\sum_{i,k=1}^{n} (d^2(A, t_i) + d^2(A, t_k) - d^2(t_i, t_k))\rho_i\rho_k \\
&\quad = \sum_{1}^{n} (r^2 + r^2 - F^2(t_i - t_k))\rho_i\rho_k \\
&\quad = \sum_{1}^{n} \left\{2r^2 - \frac{1}{2}\int_0^\infty [1 - \cos 2(t_i - t_k)u]d\alpha(u)\right\}\rho_i\rho_k \\
&\quad = \frac{1}{2}\int_0^\infty \left\{\sum_{1}^{n} \cos 2(t_i - t_k)u \cdot \rho_i\rho_k\right\} d\alpha(u) \\
&\quad = \frac{1}{2}\int_0^\infty \left\{\left(\sum_{1}^{n} \rho_i \cos 2t_iu\right)^2 + \left(\sum_{1}^{n} \rho_i \sin 2t_iu\right)^2\right\} d\alpha(u) \geqq 0.
\end{aligned}$$

Hence $F(E_1)+A$ may be imbedded in $\mathfrak{H}$ by Menger's theorem. Now (1.24) shows that $F(E_1)$ is mapped into a screw line Γ lying on the sphere of radius r whose center is the image of A.

Part II. First proof of the fundamental theorem by the theory of operators

2.1. Let f_t $(-\infty<t<\infty)$ be a screw line of $\mathfrak{H}$, and $F(t)$ the corresponding screw function. For greater generality we assume that $\mathfrak{H}$ is a real complete unitary space (not necessarily separable). As in §1.1 we have

$$\|f_t - f_s\| = F(t-s), \tag{2.1}$$

$$F(t) \to 0 \quad \text{if} \quad t \to 0. \tag{2.2}$$

Using the reality of $\mathfrak{H}$ we have

$$\begin{aligned}(f_t - f_u, f_s - f_u) &= \tfrac{1}{2}\{\|f_t - f_u\|^2 + \|f_s - f_u\|^2 - \|(f_t - f_u) - (f_s - f_u)\|^2\} \\ &= \tfrac{1}{2}\{\|f_t - f_u\|^2 + \|f_s - f_u\|^2 - \|f_t - f_s\|^2\} \\ &= \tfrac{1}{2}\{F^2(t-u) + F^2(s-u) - F^2(t-s)\}.\end{aligned}$$

Hence setting

$$F^2(p) + F^2(q) - F^2(p-q) = 2G(p, q),$$

we have

$$(f_t - f_u, f_s - f_u) = G(t-u, s-u). \tag{2.3}$$

Let $\mathfrak{H}_1$ be the closed linear subset of $\mathfrak{H}$, which is spanned by all

$$f_t - f_0, \qquad t \text{ rational.}$$

Write all rational numbers as a sequence $t_1, t_2, \cdots$, and then orthogonalize the sequence

$$f_{t_1} - f_0, f_{t_2} - f_0, \cdots$$

by the Gram–E. Schmidt procedure, thus obtaining the (finite or enumerably infinite) normalized orthogonal set

$$\phi_1, \phi_2, \cdots.$$

Thus

$$\phi_i = \sum_{j=1}^{i} \alpha_{ij}(f_{t_j} - f_0), \qquad \text{all } \alpha_{ij} \text{ are real numbers.} \tag{2.4}$$

The $f_t - f_0$, t rational, are linear aggregates of the ϕ_i's, hence (owing to f_t's continuity in t) all $f_t - f_0$, $-\infty < t < \infty$, are limit points of such linear aggregates. Therefore all $f_t - f_0$ belong to the closed, linear set which is spanned by $\phi_1, \phi_2, \cdots$, and which therefore coincides with $\mathfrak{H}_1$. Hence

$$f_t - f_0 = \sum_i \beta_i(t)\phi_i \tag{2.5}$$

where $\beta_i(t)$ are real continuous functions of t.

Combining (2.4) and (2.5) we get

$$f_t - f_0 = \sum_i \beta_i(t)\left\{\sum_{j=1}^{i} \alpha_{ij}(f_{t_j} - f_0)\right\}, \tag{2.6}$$

that is

$$\left\| (f_t - f_0) - \sum_i \beta_i(t)\left\{\sum_{j=1}^{i} \alpha_{ij}(f_{t_j} - f_0)\right\}\right\|^2 = 0. \tag{2.7}$$

Consider the equation

$$(2.6') \qquad f_{t+s} - f_s = \sum_i \beta_i(t) \left\{ \sum_{j=1}^{i} \alpha_{ij}(f_{t_j+s} - f_s) \right\},$$

and its equivalent

$$(2.7') \qquad \left\| (f_{t+s} - f_s) - \sum_i \beta_i(t) \left\{ \sum_{j=1}^{i} \alpha_{ij}(f_{t_j+s} - f_s) \right\} \right\|^2 = 0.$$

(2.7′) can be written as a relation of the

$$(f_{t'+s} - f_s, f_{t''+s} - f_s) \qquad \text{for } t', t'' = t, t_1, t_2, \cdots,$$

and the

$$\beta_i(t), \qquad \alpha_{ij}.$$

By (2.3) this means a relation of the

$$G(t', t'') \qquad \text{for } t', t'' = t, t_1, t_2, \cdots,$$

and the

$$\beta_i(t), \qquad \alpha_{ij}.$$

Hence (2.7′), and consequently (2.6′), are independent of s. But for $s=0$ (2.6′), (2.7′) coincide with (2.6), (2.7), and hence are true. Therefore they hold for all s.

Let

$$(2.4') \qquad \phi_i(s) = \sum_{j=1}^{i} \alpha_{ij}(f_{t_j+s} - f_s).$$

Then (2.6′) gives

$$(2.5') \qquad f_{t+s} - f_s = \sum_i \beta_i(t)\phi_i(s).$$

The relation

$$(2.8) \qquad (\phi_i(s), \phi_k(s)) = \delta_{ik} \begin{cases} = 1 & \text{if } i = k, \\ = 0 & \text{if } i \neq k, \end{cases}$$

means

$$\left(\sum_{j=1}^{i} \alpha_{ij}(f_{t_j+s} - f_s), \sum_{j=1}^{k} \alpha_{kj}(f_{t_j+s} - f_s) \right) = \delta_{ik},$$

which is a relation of the

$$(f_{t'+s} - f_s, f_{t''+s} - f_s), \qquad \text{for } t', t'' = t_1, t_2, \cdots,$$

and the

$$\alpha_{ij}, \qquad \alpha_{kj}.$$

By (2.3) this means a relation of

$$G(t', t''), \qquad \text{for } t', t'' = t_1, t_2, \cdots,$$

and the

$$\alpha_{ij}, \qquad \alpha_{kj}.$$

Hence (2.8) is independent of s. It expresses that $\phi_1(s), \phi_2(s), \cdots$ is a normalized orthogonal set. Since $\phi_i(0) = \phi_i$, this is true for $s = 0$. This proves that

(2.9) $\phi_1(s), \phi_2(s), \cdots$ is a normalized orthogonal set for every s.

2.2. Consider a fixed s. The $\phi_1(s), \phi_2(s), \cdots$ span the same closed, linear set as the $f_{t_1+s} - f_s, f_{t_2+s} - f_s, \cdots$, that is, as the $f_{t+s} - f_s$, t rational. Owing to the continuity of f_u in u, this is the same set as spanned by the $f_{t+s} - f_s$, $-\infty < t < +\infty$, or, if we write t for $t+s$, by the $f_t - f_s$, $-\infty < t < +\infty$.

This set contains in particular $f_0 - f_s$, hence every $f_t - f_0 = (f_t - f_s) - (f_0 - f_s)$. Hence it contains $\mathfrak{H}_1$. But $\mathfrak{H}_1$ contains all $f_t - f_0$ (by (2.5), along with the ϕ_i), hence $f_s - f_0$ and the $f_t - f_s = (f_t - f_0) - (f_s - f_0)$. Hence it contains our set, too. In other words:

(2.10) $\phi_1(s), \phi_2(s), \cdots$ span the closed linear set $\mathfrak{H}_1$ for every s.

By (2.9), (2.10) the equations

$$U(s)\phi_i(0) = \phi_i(s), \qquad \text{for } i = 1, 2, \cdots, \tag{2.11}$$

define a unitary transformation $U(s)$ in $\mathfrak{H}_1$. (2.5′) gives immediately

$$U(s)(f_t - f_0) = f_{t+s} - f_s, \tag{2.12}$$

and by (2.4′) this relation (2.12) again implies (2.11).

If we write (2.12) for $t = u$ and v, and subtract, we get

$$U(s)(f_u - f_v) = f_{u+s} - f_{v+s}. \tag{2.13}$$

Since (2.12) is a special case of (2.13) ($u = t$, $v = 0$) we see that (2.13) is also characteristic for $U(s)$. Application of (2.13) for $s = s_1$, s_2, $s_1 + s_2$ shows that

$$U(s_2)U(s_1) = U(s_1 + s_2). \tag{2.14}$$

As f_{t_j+s}, f_s are continuous functions of s, so is $\phi_i(s)$ by (2.4′). Now (2.11) shows that

(2.15) $U(s)$ is a (strongly) continuous function of s.

2.3. Extend now $\mathfrak{H}$ to a complex unitary space $\overline{\mathfrak{H}}$, and correspondingly $\mathfrak{H}_1$ to $\overline{\mathfrak{H}}_1$. Then $U(s)$ extends to a unitary operator in $\overline{\mathfrak{H}}_1$. Since the extended meaning of $\|f-g\|$ and $U(s)$ remains unchanged in $\mathfrak{H}_1$, it is clear that (2.1), (2.12) and (2.15) remain true.

As $\overline{\mathfrak{H}}_1$ is spanned by the enumerable set $f_{t_1}-f_0, f_{t_2}-f_0, \cdots$, it is separable, hence a finite-dimensional unitary or a complex Hilbert space. Now we may apply, considering (2.14), (2.15), the theorem of M. H. Stone ([6, 7]; this theorem is independent of the separability of the underlying unitary space). This theorem insures the existence of a self-adjoint operator $\overline{A}$ in $\overline{\mathfrak{H}}_1$, such that

$$U(s) = \exp(is\overline{A}). \tag{2.16}$$

Now equation (2.12) may be written as

$$f_{t+s} = U(s)(f_t - f_0) + f_s,$$

or, using (2.16),

$$f_{t+s} = \exp(is\overline{A})(f_t - f_0) + f_s. \tag{2.17}$$

2.4. Let $\overline{E}(x)$ be the resolution of unity which belongs to $\overline{A}$ in $\overline{\mathfrak{H}}_1$. Define the projections $\overline{F}_{n,\epsilon}$, $n=0, \pm 1, \pm 2, \cdots, \epsilon = \pm 1$, and $\overline{F}^0$ as follows:

$$\begin{aligned} \overline{F}_{n,+1} &= \overline{E}(2^n) - \overline{E}(2^{n-1}), \\ \overline{F}_{n,-1} &= \overline{E}(-2^{n-1}) - \overline{E}(-2^n), \\ \overline{F}^0 &= \overline{E}(0) - \overline{E}(-0) \qquad \left(E(-0) = \lim_{x<0,\, x\to 0} \overline{E}(x)\right). \end{aligned} \tag{2.18}$$

The projection operators $\overline{F}_{n,\epsilon}$, $\overline{F}^0$ are clearly orthogonal and their sum is 1.

Let

$$\begin{aligned} &\overline{\mathfrak{M}}_{n,\epsilon} \text{ be the closed linear set with the projection } \overline{F}_{n,\epsilon}, \\ &\qquad\qquad (n = 0, \pm 1, \pm 2, \cdots, \epsilon = \pm 1), \\ &\overline{\mathfrak{M}}^0 \text{ be the closed linear set with the projection } \overline{F}^0. \end{aligned} \tag{2.19}$$

Again the $\overline{\mathfrak{M}}_{n,\epsilon}$, $\overline{\mathfrak{M}}^0$ are mutually orthogonal and they span together the closed linear set $\overline{\mathfrak{H}}_1$. Clearly every $\overline{\mathfrak{M}}_{n,\epsilon}$, $\overline{\mathfrak{M}}^0$ reduces $\overline{A}$, and hence, by (2.16), all $U(s)$, too.

Denote the projections of f_t in $\overline{\mathfrak{M}}_{n,\epsilon}$ and $\overline{\mathfrak{M}}^0$, by $\overline{f}_{n,\epsilon,t}$ and $\overline{f}^0_t$, respectively.

2.5. Let us consider first the situation in $\overline{\mathfrak{M}}^0$. Here clearly $\overline{A}=0$, and (2.17) gives

$$\overline{f}^0_{t+s} = \overline{f}^0_t - \overline{f}^0_0 + \overline{f}^0_s, \tag{2.20}$$

or

$$(\overline{f}^0_{t+s} - \overline{f}^0_0) = (\overline{f}^0_t - \overline{f}^0_0) + (\overline{f}^0_s - \overline{f}^0_0).$$

This, and the t-continuity of $\overline{f}_t$, give immediately

$$\overline{f}^0_t - \overline{f}^0_0 = t(\overline{f}^0_1 - \overline{f}^0_0),$$

that is,

$$\vec{f}_t^0 = t\vec{\omega}^0 + \vec{f}_0^0, \qquad \vec{\omega}^0 \text{ being fixed.} \tag{2.21}$$

2.6. Let us consider next the situation in $\overline{\mathfrak{M}}_{n,\epsilon}$. Here the spectrum $S_{n,\epsilon}$ of $\overline{A}$ lies clearly between 2^{n-1} and 2^n, and between -2^n and -2^{n-1} respectively. Hence $x \in S_{n,\epsilon}$ implies $2^{n-1} \leqq |x| \leqq 2^n$.

Assume

$$0 < |t| \leqq \pi/2^n. \tag{2.22}$$

Then $x \in S_{n,\epsilon}$ implies $0 < 2^{n-1}|t| \leqq |tx| \leqq \pi$. Hence the function

$$f_t(x) = (\exp(itx) - 1)^{-1}$$

is everywhere defined, continuous, and bounded in $x \in S_{n,\epsilon}$. Hence we may form $f_t(\overline{A})$. As $(\exp(itx)-1) f_t(x) = f_t(x)(\exp(itx)-1) = 1$ for all $x \in S_{n,\epsilon}$, we have $(\exp(it\overline{A})-1) f_t(\overline{A}) = f_t(\overline{A})(\exp(it\overline{A})-1) = 1$. Hence $f_t(\overline{A})$ is the inverse of $\exp(it\overline{A})-1$.

We have thus proved that

$$(\exp(it\overline{A}) - 1)^{-1} \text{ exists and is bounded.} \tag{2.23}$$

Consider now (2.17) for t, s and for s, t. As the left side is the same in both cases, the right sides are equal, too, giving

$$\exp(is\overline{A})(\vec{f}_{n\epsilon/t} - \vec{f}_{n\epsilon/0}) + \vec{f}_{n\epsilon/s} = \exp(it\overline{A})(\vec{f}_{n\epsilon/s} - \vec{f}_{n\epsilon/0}) + \vec{f}_{n\epsilon/t},$$

that is

$$\{\exp(is\overline{A}) - 1\}(\vec{f}_{n\epsilon/t} - \vec{f}_{n\epsilon/0}) = \{\exp(it\overline{A}) - 1\}(\vec{f}_{n\epsilon/s} - \vec{f}_{n\epsilon/0}).$$

Assume (2.22) for both t and s. Since $\exp(it\overline{A})-1$ and $\exp(is\overline{A})-1$ commute, and by (2.23) for t and s, we may write this as

$$\{\exp(it\overline{A}) - 1\}^{-1}(\vec{f}_{n\epsilon/t} - \vec{f}_{n\epsilon/0}) = \{\exp(is\overline{A}) - 1\}^{-1}(\vec{f}_{n\epsilon/s} - \vec{f}_{n\epsilon/0}).$$

That is,

$$\{\exp(it\overline{A}) - 1\}^{-1}(\vec{f}_{n\epsilon/t} - \vec{f}_{n\epsilon/0}) = \vec{v}_{n\epsilon},$$

where $\vec{v}_{n\epsilon}$ is fixed. But this may be written as

$$\vec{f}_{n\epsilon/t} = \{\exp(it\overline{A}) - 1\}\vec{v}_{n\epsilon} + \vec{f}_{n\epsilon/0}, \qquad \vec{v}_{n\epsilon} \text{ constant.} \tag{2.24}$$

This holds provided t satisfies (2.22). Since it is obviously true for $t=0$, it holds whenever $|t| \leqq \pi/2^n$. But if (2.24) holds for both t and s then (2.17) extends it to $t+s$. Hence (2.24) holds for all t.

We may now formulate (2.21), (2.24) as follows:

(2.25) The projections of $f_t - f_0$ in $\overline{\mathfrak{M}}^0$ and $\overline{\mathfrak{M}}_{n,\epsilon}$ are $t\vec{\omega}^0$ and $\{\exp(it\overline{A}) - 1\}\vec{v}_{n\epsilon}$ respectively, the $\vec{\omega}^0$, $\vec{v}_{n\epsilon}$ being fixed elements of $\overline{\mathfrak{M}}^0$ and $\overline{\mathfrak{M}}_{n,\epsilon}$ respectively.

2.7. We know that $\overline{\mathfrak{H}}_1$ is the closed linear set spanned by the mutually orthogonal sets $\overline{\mathfrak{M}}^0$ and $\overline{\mathfrak{M}}_{n,\epsilon}$ $(n=0, \pm 1, \pm 2, \cdots, \epsilon=\pm 1)$. As $f_t - f_0$ is in $\mathfrak{H}_1$, we see that (2.25) implies

$$f_t - f_0 = t\bar{\omega}^0 + \sum_{n=-\infty}^{\infty} \sum_{\epsilon=\pm 1} \{\exp(it\overline{A}) - 1\}\bar{v}_{n\epsilon}, \tag{2.26}$$

the convergence of the right side being certain for all t.

Since the addends of the right side of (2.26) are mutually orthogonal, we conclude that

$$\|f_t - f_0\|^2 = t^2\|\bar{\omega}^0\|^2 + \sum_{n=-\infty}^{\infty} \sum_{\epsilon=\pm 1} \|\{\exp(it\overline{A}) - 1\}\bar{v}_{n\epsilon}\|^2.$$

But for each term of this sum we have

$$\begin{aligned}\|\{\exp(it\overline{A}) - 1\}\bar{v}_{n\epsilon}\|^2 &= (\{\exp(it\overline{A}) - 1\}\bar{v}_{n\epsilon}, \{\exp(it\overline{A}) - 1\}\bar{v}_{n\epsilon}) \\ &= (\{\exp(it\overline{A}) - 1\}^*\{\exp(it\overline{A}) - 1\}\bar{v}_{n\epsilon}, \bar{v}_{n\epsilon}).\end{aligned}$$

By well known properties of functions of Hermitian operators, and since

$$\begin{aligned}\overline{(\exp(itx) - 1)}(\exp(itx) - 1) &= (\exp(-itx) - 1)(\exp(itx) - 1) \\ &= 2 - \exp(itx) - \exp(-itx) = 2(1 - \cos tx),\end{aligned}$$

the above expression is equal to

$$(2(1 - \cos(it\overline{A}))\bar{v}_{n\epsilon}, \bar{v}_{n\epsilon}),$$

that is, to

$$\int 2(1 - \cos tx)d(\|\overline{E}(x)\bar{v}_{n\epsilon}\|^2).$$

It suffices to extend this Stieltjes integral from 2^{n-1} to 2^n, if $\epsilon=+1$; and from -2^n to -2^{n-1} if $\epsilon=-1$. Denote this interval by $I_{n,\epsilon}$.

We have thus obtained

$$F^2(t) = \|f_t - f_0\|^2 = t^2\|\bar{\omega}^0\|^2 + \sum_{n=-\infty}^{\infty} \sum_{\epsilon=\pm 1} \int_{I_{n,\epsilon}} 2(1 - \cos tx)d(\|\overline{E}(x)\bar{v}_{n\epsilon}\|^2). \tag{2.27}$$

Integrating this relation from 0 to 1 we obtain

$$\int_0^1 F^2(t)dt = \tfrac{1}{3}\|\bar{\omega}^0\|^2 + \sum_{n=-\infty}^{\infty} \sum_{\epsilon=\pm 1} \int_{I_{n,\epsilon}} 2\left(1 - \frac{\sin x}{x}\right)d(\|\overline{E}(x)\bar{v}_{n\epsilon}\|^2). \tag{2.28}$$

Since

$$2\left(1 - \frac{\sin x}{x}\right)\begin{cases} \geqq \tfrac{1}{4}x^2, & \text{if } |x| \leqq 1, \\ \geqq \tfrac{1}{2}, & \text{if } |x| \geqq 1, \end{cases}$$

we derive from (2.28) the finiteness of the following expression:

$$(2.28') \qquad \sum_{n=-\infty}^{0} \sum_{\epsilon=\pm 1} \int_{I_{n,\epsilon}} \tfrac{1}{4}x^2 d\left\|\overline{E}(x)\bar{v}_{n\epsilon}\right\|^2 + \sum_{n=1}^{\infty} \sum_{\epsilon=\pm 1} \int_{I_{n,\epsilon}} \tfrac{1}{4} d\left\|\overline{E}(x)\bar{v}_{n\epsilon}\right\|^2.$$

2.8. We may now readily complete a proof of Theorem 1 by showing that the last side of (2.27) can be put in the form (1.4). Indeed, define two non-decreasing functions $\beta_-(x)$ $(-\infty < x < 0)$ and $\beta_+(x)$ $(0 < x < +\infty)$ as follows: Set

$$\beta_-(x) = \left\|\overline{E}(x)\bar{v}_{0,-1}\right\|^2 \quad \text{in} \quad I_{0,-1}\ (-1 \leqq x \leqq -\tfrac{1}{2}),$$

and extend the definition successively to $I_{1,-1}$, $I_{2,-1}$, $\cdots$, and also to $I_{-1,-1}$, $I_{-2,-1}$, $\cdots$, in such a way that

(i) $\beta_-(x) = \left\|\overline{E}(x)\bar{v}_{n,-1}\right\|^2 + \text{const.}$, in $I_{n,-1}$ $(-2^n \leqq x \leqq -2^{n-1})$,

(ii) the constants are successively determined by the requirement that we have agreement of values of $\beta_-(x)$, at every point $x = -2^n$, resulting from this function's definition (i) in the adjoining intervals $I_{n,-1}$ and $I_{n+1,-1}$.

Set similarly

$$\beta_+(x) = \left\|\overline{E}(x)\bar{v}_{0,1}\right\|^2 \quad \text{in} \quad I_{0,1}\ (\tfrac{1}{2} \leqq x \leqq 1),$$

and define successively

$$\beta_+(x) = \left\|\overline{E}(x)\bar{v}_{n,1}\right\|^2 = \text{const.}, \quad \text{in} \quad I_{n,1}\ (2^{n-1} \leqq x \leqq 2^n),$$

determining the constants successively by the same requirement as above. The finite expression (2.28′) now becomes

$$(2.28'') \qquad \int_{-1}^{-0} \tfrac{1}{4}x^2 d\beta_-(x) + \int_{+0}^{1} \tfrac{1}{4}x^2 d\beta_+(x) + \int_{-\infty}^{-1} \tfrac{1}{4} d\beta_-(x) + \int_{1}^{\infty} \tfrac{1}{4} d\beta_+(x),$$

and (2.27) may be written as

$$F^2(t) = t^2\left\|\bar{\omega}^0\right\|^2 + \int_{-\infty}^{-0} 2(1 - \cos tx)d\beta_-(x) + \int_{+0}^{\infty} 2(1 - \cos tx)d\beta_+(x).$$

Setting

$$\alpha(x) = 2\beta_+(x) - 2\beta_-(-x), \qquad 0 < x < \infty,$$

this becomes

$$(2.27') \qquad F^2(t) = t^2\left\|\bar{\omega}^0\right\|^2 + \int_{+0}^{\infty} (1 - \cos tx)d\alpha(x)$$

where $\alpha(x)$ is non-decreasing such that

$$\alpha(\infty) = 2\beta_+(\infty) - 2\beta_-(-\infty),$$

$$\int_{+0}^{1} x^2 d\alpha(x) = 2\int_{+0}^{1} x^2 d\beta_+(x) + 2\int_{-1}^{-0} x^2 d\beta_-(x)$$

are finite, on account of the finiteness of (2.28''). If we introduce a new monotone function by setting

$$\gamma(0) = 0, \qquad \gamma(x) = \int_{+0}^{x} 2x^2 d\alpha(2x), \qquad x > 0,$$

we have

$$\int_{+0}^{\infty} (1 - \cos tx) d\alpha(x) = \int_{+0}^{\infty} \frac{1 - \cos 2tx}{2x^2} 2x^2 d\alpha(2x) = \int_{+0}^{\infty} \frac{\sin^2 tx}{x^2} d\gamma(x).$$

By changing, finally, the value of $\gamma(0)$ from 0 to $-\|\vec{\omega}^0\|^2$, we see that (2.27') now assumes the form (1.4). Also (1.5) is satisfied, since $\alpha(x)$ is bounded.

PART III. SECOND PROOF AND EXTENSION OF FUNDAMENTAL THEOREM BY THE THEORY OF POSITIVE DEFINITE FUNCTIONS

3.1. We know that the class of screw functions of Hilbert space is identical with the class of continuous, even, non-negative functions $F(t)$ $(F(0) = 0)$ such that the metric transform $F(E_1)$ is isometrically imbeddable in $\mathfrak{H}$ (Introduction, §1). Let us denote this class of functions by the symbol $\Pi(E_1)$. We now extend the problem of determining the class $\Pi(E_1)$, as follows: *Determine the class $\Pi(E_m)$ of functions $F(t)$ such that the metric transform $F(E_m)$ be isometrically imbeddable in $\mathfrak{H}$* $(1 \leqq m \leqq \infty, E_\infty = \mathfrak{H})$. (See [5, Introduction and §5.2].)

The definition of $\Pi(E_m)$ implies the relations

$$\Pi(E_1) \supset \Pi(E_2) \supset \cdots \supset \Pi(E_m) \supset \cdots \supset \Pi(\mathfrak{H}),$$

showing that we have to deal with an infinite sequence of non-increasing subsets of $\Pi(E_1)$. The class $\Pi(\mathfrak{H})$ was found to be identical with the class of functions $F(t)$ whose squares are of the form

$$F^2(t) = \int_0^{\infty} \frac{1 - e^{-t^2 u}}{u} d\gamma(u),$$

where $\gamma(u)$ is non-decreasing for $u \geqq 0$ and such that $\int_1^{\infty} u^{-1} d\gamma(u)$ exists [5, §5.4, Theorem 6]. We shall now determine the classes $\Pi(E_m)$, $(1 \leqq m < \infty)$.

The statement of our result requires the integral function

$$\begin{aligned} \Omega_m(t) &= 1 - \frac{t^2}{2 \cdot m} + \frac{t^4}{2 \cdot 4 \cdot m(m+2)} - \frac{t^6}{2 \cdot 4 \cdot 6 \cdot m(m+2)(m+4)} + \cdots \\ &= \Gamma\left(\frac{m}{2}\right)\left(\frac{2}{t}\right)^{(m-2)/2} J_{(m-2)/2}(t). \end{aligned} \tag{3.1}$$

In particular

$$\Omega_1(t) = \cos t, \qquad \Omega_2(t) = J_0(t), \qquad \Omega_3(t) = \sin t/t.$$

THEOREM 7. *The class* $\Pi(E_m)$ *of non-negative continuous functions* $F(t)$ $(t \geqq 0,\ F(0)=0)$ *such that the metric transform* $F(E_m)$ *be isometrically imbeddable in* $\mathfrak{H}$, *is identical with the class of functions whose squares are of the form*

$$F^2(t) = \int_0^\infty \frac{1-\Omega_m(tu)}{u^2}\, d\gamma(u), \tag{3.2}$$

where $\gamma(u)$ *is non-decreasing for* $u \geqq 0$, $\gamma(0)=0$, *and such that*

$$\int_1^\infty \frac{d\gamma(u)}{u^2} \quad \textit{exists.} \tag{3.3}$$

If $m=1$ we obtain our Theorem 1 on screw functions, since $1-\Omega_1(tu) = 1-\cos tu = 2\sin^2(tu/2)$.

The reasoning which leads to the characterization of the class $\Pi(E_1)$ given by Lemma 1, §1.1, extends, of course, to the class $\Pi(E_m)$ and gives the following statement:

A continuous non-negative function $F(t)$ $(t \geqq 0,\ F(0)=0)$ *belongs to* $\Pi(E_m)$ *if and only if it satisfies the inequality*

$$\sum_{i,k=1}^{n} \left\{F^2(P_0P_i) + F^2(P_0P_k) - F^2(P_iP_k)\right\}\rho_i\rho_k \geqq 0, \tag{3.4}$$

for any points $P_0, \cdots, P_n$ *of* E_m *and arbitrary real* ρ_i $(n=2, 3, \cdots)$, *the quantities* P_iP_k *denoting euclidean distances in* E_m.

We refer to [5, §8.1], for a simple proof, based on the statement above, that any $F(t)$ given by (3.2) belongs to $\Pi(E_m)$.

3.2. The proof of the converse, i.e., that formula (3.2) furnishes *all* elements of $\Pi(E_m)$, requires the following limit theorem:

LEMMA 2. *Let*

$$f_n(t) = \int_0^\infty \frac{1-\Omega_m(tu)}{u^2}\, d\gamma_n(u), \qquad n = 1, 2, 3, \cdots, \tag{3.5}$$

be a sequence of functions with non-decreasing $\gamma_n(u)$, $(\gamma_n(0)=0)$, *such that all integrals* $\int_1^\infty u^{-2}d\gamma_n(u)$ *exist. Let* $f_n(t)$ *converge, as* $n\to\infty$, *uniformly in any finite interval, to a function* $f(t)$. *We indicate this type of convergence by the relation*

$$f_n(t) \to f(t). \tag{3.6}$$

Then $f(t)$ *is also of the form*

$$f(t) = \int_0^\infty \frac{1-\Omega_m(tu)}{u^2}\, d\gamma(u), \quad \gamma(u)\uparrow,\ \gamma(0)=0, \quad \int_1^\infty u^{-2}d\gamma(u) \textit{ exists,} \tag{3.7}$$

and

$$\gamma_n(u) \to \gamma(u)(^2), \tag{3.8}$$

$$\int_a^\infty \frac{d\gamma_n(u)}{u^2} \to \int_a^\infty \frac{d\gamma(u)}{u^2} \quad \textit{in all continuity points } u = a > 0 \textit{ of } \gamma(u). \tag{3.9}$$

Conversely, if the relations (3.8) *and* (3.9) *hold and* $f_n(t)$, $f(t)$ *are defined by* (3.5) *and* (3.7), *then* (3.6) *holds.*

Postponing a proof of this limit theorem we shall now use it to establish Theorem 7.

Let $F(t) \in \Pi(E_m)$, that is, $F(E_m)$ be imbeddable in $\mathfrak{H}$. By [5, §5.3, Corollary 1] we conclude that the functions $\exp\{-\lambda F^2(t)\}$, $(\lambda > 0)$, belong to the class $\mathfrak{P}(E_m)$ of functions which are positive definite in E_m. By [5, §1.2, Theorem 1], we are allowed to set

$$\exp\{-\lambda F^2(t)\} = \int_0^\infty \Omega_m(tu)\,d\alpha(u, \lambda), \qquad \lambda > 0,$$

$\alpha(u, \lambda)$ being a family of non-decreasing functions, $\alpha(0, \lambda) = 0$, $\alpha(\infty, \lambda) = 1$. Now

$$\frac{1 - \exp\{-\lambda F^2(t)\}}{\lambda} = \frac{1}{\lambda}\int_0^\infty [1 - \Omega_m(tu)]\,d\alpha(u, \lambda),$$

and if we set

$$\beta(u, \lambda) = \frac{1}{\lambda}\int_0^u u^2\,d\alpha(u, \lambda),$$

we obtain

$$f(t, \lambda) = \frac{1 - \exp\{-\lambda F^2(t)\}}{\lambda} = \int_0^\infty \frac{1 - \Omega_m(tu)}{u^2}\,d\beta(u, \lambda). \tag{3.10}$$

On the other hand,

$$\lim_{\lambda \to +0} f(t, \lambda) = F^2(t)$$

holds *uniformly in any finite interval.* For, if $0 \leqq t \leqq T$, we have $F^2(t) < K$ (since $F(t)$ is continuous) and therefore

$$\begin{aligned} |\exp\{-\lambda F^2(t)\} - 1 + \lambda F^2(t)| &= |\lambda^2 F^4(t)/2! - \lambda^3 F^6(t)/3! + \cdots| \\ &\leqq \lambda^2 K^2 e^{\lambda K}, \end{aligned}$$

or

(2) Here and throughout this paper a limiting relation $\gamma_n(u) \to \gamma(u)$ involving monotone functions is assumed to hold for all values of u which are continuity points of the limiting function $\gamma(u)$.

$$\left|\frac{1-\exp\{-\lambda F^2(t)\}}{\lambda}-F^2(t)\right|\leqq \lambda K^2 e^{\lambda K},$$

a bound which is independent of t and tends to zero with λ. Setting $f_n(t)=f(t, 1/n)$ we therefore have $f_n(t)\to F^2(t)$ and Lemma 2 is applicable, on account of (3.10), showing that $F^2(t)$ is of the form (3.2).

3.3. In this proof of Theorem 7 only a part of Lemma 2 was used, namely that (3.5) and (3.6) imply (3.7). We have stated and shall prove the more inclusive limit theorem for two reasons. First, our limit theorem implies that $\gamma(u)$ of (3.2) is uniquely defined by $F(t)$. Second, a proof of the complete statement is hardly more complicated than a proof of the partial statement actually used above.

Lemma 2 will be a direct consequence of the following analogous limit theorem involving a simpler kernel under the integral sign.

LEMMA 3. *Let*

$$\phi_n(t)=\int_0^\infty \frac{1-e^{-ut}}{u}\,d\beta_n(u), \qquad n=1, 2, 3, \cdots, \tag{3.11}$$

be a sequence of functions with non-decreasing $\beta_n(u)$ ($\beta_n(0)=0$), such that all integrals $\int_1^\infty u^{-1}d\beta_n(u)$ exist. Let $\phi_n(t)$ converge, as $n\to\infty$, to a function $\phi(t)$, uniformly in any finite interval $0\leqq t\leqq t_0$, i.e.,

$$\phi_n(t)\to\phi(t). \tag{3.12}$$

Then $\phi(t)$ is also of the form

$$\phi(t)=\int_0^\infty \frac{1-e^{-ut}}{u}\,d\beta(u),\ \beta(u)\uparrow,\ \beta(0)=0,\ \int_1^\infty u^{-1}d\beta(u) \text{ exists,} \tag{3.13}$$

and

$$\beta_n(u)\to\beta(u), \tag{3.14}$$

$$\int_a^\infty \frac{d\beta_n(u)}{u}\to\int_a^\infty \frac{d\beta(u)}{u} \text{ in all continuity points } u=a>0 \text{ of } \beta(u). \tag{3.15}$$

Conversely, if the relations (3.14) *and* (3.15) *hold and $\phi_n(t)$, $\phi(t)$ are defined by* (3.11) *and* (3.13), *then* (3.12) *holds*[3].

The $\phi_n(t)$ belong to the class T defined in [5, §4]. It was shown there (Lemma 5) that (3.12) implies $\phi(t)\in T$, hence (3.13). It was furthermore

---

[3] That the condition (3.15) cannot be dispensed with is shown by the example of $\beta_n(u)=0$ if $0\leqq u<n$, $\beta_n(u)=n$ if $u\geqq n$, $\beta(u)=0$, hence $\phi_n(t)=1-e^{-nt}$, $\phi(t)=0$. Now (3.14) holds but not (3.12). Note that the condition (3.15) is not satisfied.

shown by means of the Vitali-Porter convergence theorem that $\phi_n(t)$ and

$$\phi_n'(t) = \int_0^\infty \epsilon^{-ut} d\beta_n(u)$$

converge to $\phi(t)$ and $\phi'(t)$, respectively, uniformly in any finite closed domain inside the half-plane $\Re t > 0$ of the complex variable $t = \sigma + i\tau$. In particular we therefore have

$$(3.16) \quad \phi_n'(1 + i\tau) = \int_0^\infty \epsilon^{-u}\epsilon^{-iu\tau} d\beta_n(u) \to \phi'(1 + i\tau) = \int_0^\infty \epsilon^{-u}\epsilon^{-iu\tau} d\beta(u),$$

uniformly in any finite τ-interval. In terms of the new monotone and bounded functions

$$(3.17) \qquad \alpha_n(u) = \int_0^u \epsilon^{-u} d\beta_n(u), \qquad \alpha(u) = \int_0^u \epsilon^{-u} d\beta(u),$$

the relation (3.16) becomes

$$\int_0^\infty \epsilon^{-iu\tau} d\alpha_n(u) \to \int_0^\infty \epsilon^{-iu\tau} d\alpha(u).$$

Now we may apply a limit theorem of Paul Lévy and conclude that

$$(3.18) \qquad \alpha_n(u) \to \alpha(u),$$

[1, p. 197]. But (3.17) may be inverted:

$$\beta_n(u) = \int_0^u e^u d\alpha_n(u), \qquad \beta(u) = \int_0^u e^u d\alpha(u).$$

Now (3.18) implies the desired relation (3.14).

There remains the proving of (3.15). On one hand (3.14) implies

$$\int_0^a \frac{1 - \epsilon^{-iu}}{u} d\beta_n(u) \to \int_0^a \frac{1 - \epsilon^{-iu}}{u} d\beta(u), \qquad t \geqq 0,$$

and therefore (3.11), (3.13) and (3.12) imply

$$(3.19) \qquad \int_a^\infty \frac{1 - \epsilon^{-iu}}{u} d\beta_n(u) \to \int_a^\infty \frac{1 - \epsilon^{-iu}}{u} d\beta(u), \qquad t \geqq 0.$$

On the other hand, the obvious inequalities

$$(1 - \epsilon^{-ta}) \int_a^\infty \frac{d\beta_n(u)}{u} \leqq \int_a^\infty \frac{1 - \epsilon^{-iu}}{u} d\beta_n(u) \leqq \int_a^\infty \frac{d\beta_n(u)}{u}, \quad t \geqq 0,$$

imply

$$\int_a^\infty \frac{1-e^{-tu}}{u}\,d\beta_n(u) \leqq \int_a^\infty \frac{d\beta_n(u)}{u} \leqq \frac{1}{1-e^{-ta}}\int_a^\infty \frac{1-e^{-tu}}{u}\,d\beta_n(u), t>0.$$

As $n\to\infty$ we get by (3.19)

$$\int_a^\infty \frac{1-e^{-tu}}{u}\,d\beta(u) \leqq \liminf \int_a^\infty \frac{d\beta_n(u)}{u} \leqq \limsup \int_a^\infty \frac{d\beta_n(u)}{u}$$

$$\leqq \frac{1}{1-e^{-ta}}\cdot\int_a^\infty \frac{1-e^{-tu}}{u}\,d\beta(u).$$

As $t\to\infty$ this gives

$$\int_a^\infty \frac{d\beta(u)}{u} \leqq \liminf \int_a^\infty \frac{d\beta_n(u)}{u} \leqq \limsup \int_a^\infty \frac{d\beta_n(u)}{u} \leqq \int_a^\infty \frac{d\beta(u)}{u},$$

which proves (3.15).

In order to prove the converse of Lemma 3, we have to show that (3.11), (3.13), (3.14), and (3.15) imply (3.12). For a given ϵ, we can choose a positive a, continuity point of $\beta(u)$, such that

$$\int_a^\infty u^{-1}d\beta_n(u) < \epsilon, \qquad \int_a^\infty u^{-1}d\beta(u) < \epsilon,$$

provided n is sufficiently large. But then

$$\int_0^a \frac{1-e^{-tu}}{u}\,d\beta_n(u), \qquad \int_0^a \frac{1-e^{-tu}}{u}\,d\beta(u),$$

differ from $\phi_n(t)$ and $\phi(t)$ respectively by less than ϵ. Hence

$$|\phi_n(t)-\phi(t)| < 2\epsilon + \left|\int_0^a \frac{1-e^{-tu}}{u}\,d\beta_n(u) - \int_0^a \frac{1-e^{-tu}}{u}\,d\beta(u)\right|,$$

provided n is sufficiently large. On the other hand, (3.14) implies

$$\int_0^a \frac{1-e^{-tu}}{u}\,d\beta_n(u) \to \int_0^a \frac{1-e^{-tu}}{u}\,d\beta(u),$$

that is,

$$\left|\int_0^a \frac{1-e^{-tu}}{u}\,d\beta_n(u) - \int_0^a \frac{1-e^{-tu}}{u}\,d\beta(u)\right| < \epsilon,$$

in an arbitrarily given interval $0\leqq t\leqq t_0$, provided n is large enough. Hence $|\phi_n(t)-\phi(t)|<3\epsilon$, in the interval $0\leqq t\leqq t_0$, provided n is large enough. This is precisely identical to (3.12).

By a similar argument we can prove the converse part of Lemma 2. We

may, therefore, consider the converse part of Lemma 2 as already established and shall use it in proving the direct part of that lemma.

3.4. In order to prepare a proof of Lemma 2 we need the following statement:

(3.20) The assumptions (3.5) and (3.6), of Lemma 2, imply the existence of two constants K and H, independent of n and t, such that $f_n(t) < Kt^2 + H$, $t \geqq 0$, $n = 1, 2, 3, \cdots$.

Assume first that $m \geqq 2$. Since

$$-1 < \Omega_m(t) < 1, \quad (t > 0); \qquad \lim_{t \to \infty} \Omega_m(t) = 0,$$

we see that (3.5) and (3.6) imply (for $t = 1$) that the two expressions

$$\int_0^1 \frac{1 - \Omega_m(u)}{u^2}\, d\gamma_n(u), \qquad \int_1^\infty [1 - \Omega_m(u)]\, \frac{d\gamma_n(u)}{u^2}$$

are both bounded, as $n \to \infty$. Since the integrands $[1 - \Omega_m(u)]/u^2$, and $1 - \Omega_m(u)$, have positive lower bounds in their respective intervals of integration $(0, 1)$ and $(1, \infty)$, we see that

$$\int_0^1 d\gamma_n(u) = \gamma_n(1), \qquad \int_1^\infty \frac{d\gamma_n(u)}{u^2} \tag{3.21}$$

also are bounded, as $n \to \infty$. On the other hand, it is clear from the power series expansion (3.1), that there is a constant c, such that

$$1 - \Omega_m(t) \leqq ct^2,$$

provided $t \geqq 0$. Hence, by (3.21),

$$\begin{aligned} f_n(t) &= \int_0^1 \frac{1 - \Omega_m(tu)}{u^2}\, d\gamma_n(u) + \int_1^\infty \frac{1 - \Omega_m(tu)}{u^2}\, d\gamma_n(u) \\ &\leqq \int_0^1 \frac{ct^2u^2}{u^2}\, d\gamma_n(u) + \int_1^\infty \frac{2}{u^2}\, d\gamma_n(u) \\ &= t^2 c\gamma_n(1) + 2\int_1^\infty u^{-2} d\gamma_n(u) < Kt^2 + H. \end{aligned}$$

If $m = 1$, hence $\Omega_m(tu) = \cos tu$, we first integrate both sides of (3.6), between 0 and 1, obtaining the relation

$$\int_0^1 f_n(t)dt = \int_0^\infty \frac{1 - u^{-1}\sin u}{u^2}\, d\gamma_n(u) = \int_0^\infty \frac{1 - \Omega_3(u)}{u^2}\, d\gamma_n(u) \to \int_0^1 f(t)dt.$$

From this we derive as above that (3.21) are also bounded in the present case. The argument is then concluded as above. Thus (3.20) is established.

3.5. We may now readily prove Lemma 2. From its assumptions (3.5), (3.6), combined with (3.20), we get

$$f_n(t) \to f(t), \qquad 0 \leqq f_n(t) < Kt^2 + H, \qquad 0 \leqq f(t) < Kt^2 + H, t \geqq 0.$$

This implies, as readily seen, that

$$\int_0^\infty e^{-t^2/4} f_n(t\mu) t^{m-1} dt \to \int_0^\infty e^{-t^2/4} f(t\mu) t^{m-1} dt, \tag{3.22}$$

uniformly in μ in any finite interval $0 \leqq \mu \leqq M$. Hence also

$$\int_0^\infty e^{-t^2/4} f_n(t\lambda^{1/2}) t^{m-1} dt \to \int_0^\infty e^{-t^2/4} f(t\lambda^{1/2}) t^{m-1} dt,$$

uniformly in any finite interval $0 \leqq \lambda \leqq L$. Taking $t\lambda^{-1/2}$, instead of t, as a new variable of integration, we see that

$$\lambda^{-m/2} \int_0^\infty e^{-t^2/4} f_n(t) t^{m-1} dt \to \lambda^{-m/2} \int_0^\infty e^{-t^2/4} f(t) t^{m-1} dt, \tag{3.23}$$

uniformly in λ in any finite interval $0 < \lambda \leqq L$. We shall now make use of the relation [5, p. 823]

$$e^{-\lambda u^2} = c_m \lambda^{-m/2} \int_0^\infty \Omega_m(tu) e^{-t^2/4} t^{m-1} dt, \quad \lambda > 0, \; c_m = 2^{1-m} \cdot [\Gamma(m/2)]^{-1}, \tag{3.24}$$

and its particular case ($u = 0$)

$$1 = c_m \lambda^{-m/2} \int_0^\infty e^{-t^2/4} t^{m-1} dt.$$

If we multiply the relation

$$f_n(t) = \int_0^\infty \frac{1 - \Omega_m(tu)}{u^2} d\gamma_n(u)$$

by $c_m \lambda^{-m/2} e^{-t^2/4} t^{m-1} dt$, we get by integration

$$\phi_n(\lambda) = c_m \lambda^{-m/2} \int_0^\infty e^{-t^2/4} f_n(t) t^{m-1} dt = \int_0^\infty \frac{1 - e^{-\lambda u^2}}{u^2} d\gamma_n(u)$$

$$= \int_0^\infty \frac{1 - e^{-\lambda u}}{u} d\gamma_n(u^{1/2}).$$

By (3.23) $\phi_n(\lambda)$ converges, as $n \to \infty$, uniformly in any interval $0 < \lambda \leqq L$. As the last expression is defined and continuous for $\lambda = 0$, we see that the uniform convergence extends to any closed interval $0 \leqq \lambda \leqq L$. But then, by Lemma 3,

we conclude that there is a monotone function $\gamma(u^{1/2})$, $(u \geqq 0)$, such that

$$\gamma_n(u^{1/2}) \to \gamma(u^{1/2}), \quad \int_{a^2}^{\infty} \frac{d\gamma_n(u^{1/2})}{u} \to \int_{a^2}^{\infty} \frac{d\gamma(u^{1/2})}{u},$$

that is,

$$\gamma_n(u) \to \gamma(u), \quad \int_{a^2}^{\infty} \frac{d\gamma_n(u)}{u^2} \to \int_{a^2}^{\infty} \frac{d\gamma(u)}{u^2}. \tag{3.25}$$

By the converse of Lemma 2, we may now conclude that

$$f_n(t) = \int_0^{\infty} \frac{1 - \Omega_m(tu)}{u^2}\, d\gamma_n(u) \to \int_0^{\infty} \frac{1 - \Omega_m(tu)}{u^2}\, d\gamma(u).$$

On comparing this relation with the assumption (3.6), we derive the validity of the desired conclusion (3.7).

Bibliography

1. P. Lévy, *Calcul des Probabilités*, Paris, 1925.

2. K. Menger, *Die Metrik des Hilbertschen Raumes*, Anzeiger der Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschaftliche Klasse, vol. 65 (1928), pp. 159–160.

3. I. J. Schoenberg, *Remarks to Maurice Fréchet's article · · ·* , Annals of Mathematics, (2), vol. 36 (1935), pp. 724–732.

4. ———, *Metric spaces and positive definite functions*, these Transactions, vol. 44 (1938), pp. 522–536.

5. ———, *Metric spaces and completely monotone functions*, Annals of Mathematics, (2), vol. 39 (1938), pp. 811–841.

6. M. H. Stone, *Linear transformations in Hilbert space* III, Proceedings of the National Academy of Sciences, vol. 16 (1930), pp. 172–175.

7. ———, *On one-parameter unitary groups in Hilbert space*, Annals of Mathematics, (2), vol. 33 (1932), pp. 643–648.

8. W. A. Wilson, *On certain types of continuous transformations of metric spaces*, American Journal of Mathematics, vol. 57 (1935), pp. 62–68.

Institute for Advanced Study,
Princeton, N. J.,
Colby College,
Waterville, Me.

PROPERTIES OF GROWTH FOR SOLUTIONS OF DIFFERENTIAL EQUATIONS OF DYNAMICAL TYPE

BY

W. J. TRJITZINSKY

1. Introduction. *In this work we shall investigate properties of "growth" of solutions of differential systems, with particular reference to systems of the type of importance in dynamics.* The systems of the latter kind are nonlinear. A considerable part of the developments for the nonlinear problem can be based on the theory of the problem obtained by retaining the linear part, only, of the nonlinear system in question. In §§3 and 4 we shall give developments for linear systems. This treatment is followed by consideration of nonlinear systems of a very general type, given in §§5, 6. It is not our main purpose, at this time, to investigate problems of dynamical stability and instability. However, it is to be noted that Theorem 6.1 presents an extended case of conditional stability.

Of importance from our present point of view are a number of contributions, of which we shall mention certain memoirs due to *A. Liapounoff*(1), *P. Bohl*(2), *É. Cotton*(3), *O. Perron*(4).

In §2 we investigate growth of functions on the basis of "characteristic numbers," defined with respect to any "admissible" function (cf. Definitions 2.1, 2.3). The developments for linear differential systems are given with the aid of characteristic numbers (§3) and "product integration" (§4). In the study of nonlinear systems use is made of characteristic numbers (§5) and of the method of successive approximations (§6).

Presented to the Society, September 11, 1940; received by the editors April 1, 1940, and, in revised form, September 25, 1940.

(1) A. Liapounoff, *Problème général de la stabilité du mouvement* (translated from Russian by E. Davaux), Annales de la Faculté des Sciences de Toulouse (2), vol. 9 (1907), pp. 203–474, originally published by the Société Mathématique de Kharkow in 1892. *This work will be referred to as* (L).

(2) P. Bohl, *Sur certaines équations différentielles d'un type général utilisable en mécanique* (translated from Russian by Mlle Tarnarider), Bulletin de la Société Mathématique de France, vol. 38 (1910), pp. 5–138; also *Ueber die Bewegung eines mechanischen Systems in der Nähe einer Gleichgewichtslage*, Journal für die reine und angewandte Mathematik, vol. 127 (1904), pp. 179–276.

(3) É. Cotton, *Sur les solutions asymptotiques des équations différentielles*, Annales Scientifique de l'École Normale Supérieure, vol. 28 (1911), pp. 473–521.

(4) O. Perron, *Ueber Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssystemen*, Mathematische Zeitschrift, vol. 29 (1928), pp. 129–160; also *Die Stabilitätsfrage bei Differentialgleichungen*, ibid., vol. 32 (1930), pp. 703–728.

The main results are embodied in Theorems 2.1, 2.2, 3.1, 3.2, 4.1, 4.2, 4.3, 5.1, 5.2, 6.1. There is an important application given subsequent to Theorem 6.1.

2. **Growth of functions.** In the sequel, unless the contrary is stated, all the functions and numbers will be supposed to be real, *the functions being defined on the interval* $(t_0, +\infty)$.

It will be convenient to introduce the following definitions.

DEFINITION 2.1. *A function* $\psi(t)$, *continuous for* $t \geqq t_0$, *monotone non-decreasing for* $t \geqq t_0$, *and such that*

$$\lim \psi(t) = +\infty \qquad (\text{for } t \to +\infty)$$

will be said to be "admissible."

DEFINITION 2.2. *A function* $\gamma(t)$ *will be said to be identically zero, provided there exists a number* t_1 $(\geqq t_0)$ *so that* $\gamma(t) = 0$ *for all* $t \geqq t_1$.

DEFINITION 2.3. *A function* $x(t)$ *will be said to have a "characteristic number"* λ, *with respect to an admissible function* $\psi(t)$, *provided*

$$\limsup \left| x(t)\psi^{\lambda+\epsilon}(t) \right| = \infty, \qquad \lim \left| x(t)\psi^{\lambda-\epsilon}(t) \right| = 0 \tag{2.1}$$

$(as\ t \to +\infty)$ *for all* $\epsilon > 0$. *We shall employ the following notation for this number:*

$$\lambda = C(x(t) \mid \psi(t)). \tag{2.2}$$

When in the first relation (2.1) *one may replace* lim sup *by* lim, *it will be said that* λ *is a characteristic number in the strong sense; this number will then be designated by*

$$\lambda = C_s(x(t) \mid \psi(t)).$$

It is to be noted that characteristic numbers in the sense of (2.2), with $\psi(t) = \exp t$, have been previously used by Liapounoff (cf. (L)).

DEFINITION 2.4. *Given a function* $f(t)$, *form the limits*

$$U(t_1) = \limsup \left| f(t) \right|, \qquad u(t_1) = \liminf \left| f(t) \right| \qquad (\text{for } t \geqq t_1),$$

where t_1 *is any number* $\geqq t_0$. *We shall write*

$$\overline{L} \left| f(t) \right| = \lim U(t_1), \qquad \underline{L} \left| f(t) \right| = \lim u(t_1) \qquad (as\ t_1 \to +\infty).$$

Clearly, $U(t_1)$ is monotone non-increasing and $u(t_1)$ is monotone non-decreasing.

The advantage in the use of $\overline{L}$, $\underline{L}$ in place of lim sup, lim inf lies in the fact that the first two symbols are more descriptive of the behavior of $f(t)$ in the neighborhood of $t = +\infty$.

We shall now prove the following theorem.

THEOREM 2.1. *Let $x(t)$ be continuous for $t \geqq t_0$ and be not identically zero. Then $x(t)$ has a finite characteristic number with respect to some admissible function $\psi(t)$.*

Suppose first the $|x(t)|$ is not uniformly bounded. One then has

$$\limsup |x(t)| = \overline{L}|x(t)| = \infty. \tag{2.3}$$

There exists a non-empty class K of functions $q(t)$, continuous, monotone non-decreasing and such that

$$|x(t)| \leqq q(t).$$

In the class K there is a "least" function $\psi(t)$; that is, for every t ($\geqq t_0$) the value $\psi(t)$ is the least value which the functions of K may assume. Clearly, in view of (2.3) $\psi(t)$ is admissible and

$$|x(t)| \leqq \psi(t); \tag{2.4}$$

moreover, there exists a set of values

$$t_0 < t_1 < \cdots < t_\nu < \cdots,$$

with $\lim_\nu t_\nu = \infty$, so that

$$\psi(t_\nu) = |x(t_\nu)| \qquad (\nu = 1, 2, \cdots).$$

One has

$$l_\epsilon(t) = |x(t)\psi^{-1+\epsilon}(t)| = |x(t)\psi^{-1}(t)|\,\psi^\epsilon(t);$$

hence

$$l_\epsilon(t_\nu) = \psi^\epsilon(t_\nu) \to \infty \qquad (\text{all } \epsilon > 0;\ \text{as } \nu \to \infty),$$

$$\limsup l_\epsilon(t) = \infty \qquad (\text{all } \epsilon > 0).$$

On the other hand, by (2.4)

$$l_{-\epsilon}(t) = |x(t)\psi^{-1-\epsilon}(t)| \leqq \psi^{-\epsilon}(t).$$

Whence

$$\lim l_{-\epsilon}(t) = 0 \qquad (\text{all } \epsilon > 0).$$

Thus, when $|x(t)|$ is unbounded, one has

$$C(x(t)\,|\,\psi(t)) = -1,$$

where $\psi(t)$ is the function referred to in (2.4).

The alternative to the hypothesis implied in (2.3) is

$$|x(t)| \leqq \beta, \qquad (\text{finite } \beta > 0). \tag{2.5}$$

One then may have, when t_1 ($\geqq t_0$) is taken sufficiently great,

$$0 < \alpha \leqq |x(t)| \leqq \beta \qquad (t \geqq t_1;\ \text{constant } \alpha); \tag{2.6}$$

then, on letting $\psi(t)$ denote any admissible function, it is inferred that

$$|x(t)\psi^{\epsilon}(t)| \geqq \alpha\psi^{\epsilon}(t), \qquad |x(t)\psi^{-\epsilon}(t)| \leqq \beta\psi^{-\epsilon}(t),$$

which implies

$$\lim |x(t)\psi^{\epsilon}(t)| = \infty, \qquad \lim |x(t)\psi^{-\epsilon}(t)| = 0$$

for all $\epsilon > 0$. Whence, under (2.6),

$$C_s(x(t) \mid \psi(t)) = 0$$

for all admissible functions $\psi(t)$.

The remaining subcase of (2.5) is

$$|x(t)| \leqq \beta, \qquad \underline{L}\,|x(t)| = 0.$$

The latter relation will imply existence of values

$$(t_0 <)\ t_1' < t_2' < \cdots < t_\nu' < \cdots; \quad \lim_\nu t_\nu' = \infty$$

for which

$$\lim_\nu |x(t_\nu')| = 0.$$

This follows from the fact that the equality $\underline{L}\,|x(t)| = 0$ signifies that the function

$$u(T) = \liminf |x(t)| \qquad (t \geqq T)$$

is zero for all $T \geqq t_0$. In consequence of the continuity of $x(t)$ and inasmuch as $x(t)$ is not identically zero one may choose the t_ν' so that

$$|x(t_\nu')| > 0 \qquad (\nu = 1, 2, \cdots).$$

Amongst the (t_ν') one may select an infinite subsequence (t_ν) $(\nu = 1, 2, \cdots)$ so that

$$|x(t_1)| \geqq |x(t_2)| \geqq \cdots \geqq |x(t_\nu)| \geqq \cdots; \tag{2.7}$$

necessarily

$$|x(t_\nu)| > 0, \qquad \lim_\nu |x(t_\nu)| = 0. \tag{2.7a}$$

Let $h(u)$ be any admissible function, defined for

$$u \geqq |x(t_1)|^{-1} = u_0$$

and such that

$$\lim_\nu |x(t_\nu)h^{\epsilon}(|x(t_\nu)|^{-1})| = \infty \qquad (\text{all } \epsilon > 0). \tag{2.8}$$

For example, $h(u) = \exp u$ is such a function. In consequence of (2.7)

$$\gamma_\nu = h(|x(t_\nu)|^{-1}) \leqq \gamma_{\nu+1} \qquad (\nu = 1, 2, \cdots);$$

moreover, by (2.7a) $\lim \gamma_\nu = \infty$. Hence it is possible to construct an admissible function $\psi(t)$ for which

$$\psi(t_\nu) = \gamma_\nu \qquad (\nu = 1, 2, \cdots).$$

On writing

$$\lambda_\epsilon(t) = |x(t)\psi^\epsilon(t)|, \tag{2.9}$$

it is observed that

$$\lambda_\epsilon(t_\nu) = |x(t_\nu)h^\epsilon(|x(t_\nu)|^{-1}).$$

Whence in consequence of (2.8)

$$\lim_\nu \lambda_\epsilon(t_\nu) = \infty \qquad (\text{all } \epsilon > 0).$$

Together with (2.9), this implies that

$$\limsup |x(t)\psi^\epsilon(t)| = \infty$$

for the admissible function $\psi(t)$ introduced above.

On the other hand, on noting the relations

$$|x(t)| \leqq \beta, \qquad \underline{L}|x(t)| = 0,$$

it is observed that the function of (2.9) satisfies

$$\lambda_{-\epsilon}(t) \leqq \beta\psi^{-\epsilon}(t);$$

thus

$$\lim \lambda_{-\epsilon}(t) = 0 \qquad (\text{all } \epsilon > 0).$$

Hence, for the subcase under consideration, we have

$$C(x(t) \mid \psi(t)) = 0,$$

where $\psi(t)$ is an admissible function with properties as described above.

This establishes Theorem 2.1.

THEOREM 2.2. *Let $x(t)$ be continuous and be not identically zero. In order that*

$$C(x(t) \mid \psi(t)) = 0 \tag{2.10}$$

for all admissible functions $\psi(t)$ it is necessary and sufficient that

$$\overline{L}|x(t)| \neq 0, \qquad \neq \infty. \tag{2.11}$$

To prove that (2.10) implies (2.11) suppose, if possible, that

$$\overline{L}|x(t)| = \infty. \tag{2.12}$$

There exists then an admissible function $q(t)$ (>0) such that

$$|x(t)| \leqq q(t), \qquad q(t_\nu) = |x(t_\nu)| \to \infty \qquad (\text{as } \nu \to \infty),$$

where the t_ν are certain numbers satisfying

$$t_0 < t_1 < \cdots; \qquad \lim_\nu t_\nu = \infty.$$

Let $\phi(t)$ be any admissible function for which

$$\lim q(t)\psi^{-\epsilon}(t) = \infty \qquad (\text{all } \epsilon > 0; \text{ as } t \to \infty).$$

One may take $\phi(t) = q(t)$ and $0 < \epsilon < 1$. We write

$$\lambda_{-\epsilon}(t) = |x(t)\phi^{-\epsilon}(t)| = \left|\frac{x(t)}{q(t)}\right| q(t)\phi^{-\epsilon}(t)$$

and note that in view of the preceding

$$\left|\frac{x(t_\nu)}{q(t_\nu)}\right| = 1.$$

Whence

$$\lambda_{-\epsilon}(t_\nu) = q(t_\nu)\phi^{-\epsilon}(t_\nu) \to \infty \qquad (\text{as } \nu \to \infty)$$

for all $\epsilon > 0$ and

$$\limsup \lambda_{-\epsilon}(t) = \infty$$

for any function $\phi(t)$ of the above description. The latter relation is contrary to the hypothesis that (2.10) holds for all admissible functions $\psi(t)$. Thus (2.12) is impossible and we have

$$|x(t)| \leqq \beta < \infty. \tag{2.13}$$

Still assuming (2.10) it is now observed that failure of (2.11) is conceivable only when (2.13) holds, while

$$\overline{L}|x(t)| = 0. \tag{2.13a}$$

Now the latter relation implies that $\lim x(t) = 0$ (as $t \to \infty$). Amongst the continuous monotone non-increasing functions $c(t)$, such that

$$|x(t)| \leqq c(t),$$

there is a least function $h(t)$ with the same properties. We have

$$h(t) = \limsup |x(\tau)| \qquad (\tau \geqq t).$$

Since

$$|x(t)| \leqq h(t)$$

and since $x(t)$ is not identically zero, it is deduced that $h(t)>0$; moreover,

$$\lim_{t\to\infty} h(t) = 0.$$

It is observed that for some admissible $\phi(t)$ we have

$$\lambda_\epsilon(t) = \left| x(t)\phi^\epsilon(t) \right| \to 0 \qquad (\text{all } \epsilon > 0;\ \text{as } t \to \infty).$$

In fact, one may take

$$\phi(t) = -\log h(t).$$

Whence (2.10) is impossible for some admissible functions and (2.13a) cannot hold. Accordingly, (2.10) implies (2.11).

To establish the converse we assume (2.11) and proceed to deduce (2.10). The second relation (2.11) implies that

$$\left| x(t) \right| \leqq \beta < \infty.$$

Hence, with $\psi(t)$ denoting any admissible function, one has

$$\lambda_{-\epsilon}(t) = \left| x(t)\psi^{-\epsilon}(t) \right| \leqq \beta\psi^{-\epsilon}(t)$$

and

$$\lim \lambda_{-\epsilon}(t) = 0 \qquad (\text{all } \epsilon > 0).$$

On the other hand, on writing

$$\overline{L}\left| x(t) \right| = \alpha\ (>0)$$

and on letting α_0 denote any constant such that $0<\alpha_0<\alpha$, it is deduced that there exist values

$$t_1 < t_2 < \cdots < t_\nu < \cdots \qquad (t_0 < t_1)$$

such that

$$\lim_\nu t_\nu = +\infty, \qquad \left| x(t_\nu) \right| \geqq \alpha_0 \qquad (\nu = 1, 2, \cdots).$$

This follows from the continuity of $x(t)$. Let $\psi(t)$ again be any admissible function. By the above relations

$$\lambda_\epsilon(t_\nu) \geqq \alpha_0\psi^\epsilon(t_\nu).$$

The last member, here, tends to $+\infty$ with t, for all $\epsilon>0$. Thus

$$\limsup_{t\to\infty} \lambda_\epsilon(t) = \infty \qquad (\text{all } \epsilon > 0).$$

Together with a previous relation for $\lambda_{-\epsilon}(t)$ this establishes (2.10), which completes the proof of Theorem 2.2.

We shall define the *characteristic number* (with respect to an admissible

function $\psi(t)$ *of a finite set of functions* $x_j(t)$ $(j=1, \cdots, m)$ as the least one of the characteristic numbers of the $x_j(t)$. This is in agreement with Liapounoff's definition for the case of $\psi(t)=\exp t$. Our designation for the characteristic number (with respect to $\psi(t)$ of functions $x_1(t), \cdots, x_m(t)$ will be

$$C(x_1, \cdots, x_m(t) \mid \psi(t)).$$

3. Characteristic numbers for linear systems. Consider the linear differential system

$$\frac{dx_i}{dt} = p_{1,i}(t)x_1 + \cdots + p_{n,i}(t)x_n \qquad (i = 1, \cdots, n), \tag{3.1}$$

where the $p_{i,j}(t)$ are continuous for $t \geqq t_0$.

Let $p(t)$ be the "least" function such that

$$| p_{i,j}(t) | \leqq p(t) \qquad (i, j = 1, \cdots, n). \tag{3.2}$$

Necessarily $p(t)$ is continuous. Inasmuch as $p(t) \geqq 0$, the function

$$\phi(t) = \int_{t_0}^{t} p(t)dt \tag{3.2a}$$

will be continuous, monotone non-decreasing.

The transformation

$$\tau = \phi(t), \tag{3.3}$$

applied to (3.1), will yield

$$\frac{dx_i}{d\tau} = h_{1,i}(\tau)x_1 + \cdots + h_{n,i}(\tau)x_n \qquad (i = 1, \cdots, n),$$

where

$$h_{i,j}(\tau) = \frac{p_{i,j}(t)}{p(t)}.$$

In view of (3.3) and (3.2a) to the interval $(t_0 \leqq t < \infty)$ corresponds the interval $(0 \leqq \tau \leqq \tau')$, with

$$\tau' = \int_{t_0}^{\infty} p(t)dt,$$

the convention being made that when the latter integral diverges one has $\tau' = \infty$. In any case, by virtue of (3.2),

$$| h_{i,j}(\tau) | \leqq 1 \qquad (0 \leqq \tau \leqq \tau').$$

LEMMA 3.1. *Every linear differential system* (3.1) *is transformed by* (3.3) [(3.2a), (3.2)] *into a system of the same type with coefficients uniformly bounded*

in absolute value on an interval which is infinite or finite, according as the integral

$$\int_{t_0}^{\infty} p(t)dt \tag{3.4}$$

diverges or converges.

We recall now a certain theorem, due to *Liapounoff*(8). In the notation at the end of §2, the theorem states that for every solution

$$x_1(t), \cdots, x_n(t)$$

of a system (3.1), whose coefficients are continuous and uniformly bounded in absolute value for $t \geqq t_0$, the number

$$C(x_1(t), \cdots, x_n(t) \mid e^t) \tag{3.5}$$

is finite.

This fact, together with Lemma 3.1, enables us at once to make the following inference with respect to characteristic numbers of solutions of *any* system (3.1).

We form the function $p(t)$, as stated in connection with (3.2). *If the integral* (3.4) *converges*, then $p(t)$ is uniformly bounded for $t \geqq t_0$; the same will be true of the $|p_{i,j}(t)|$ $(i, j = 1, \cdots, n)$ and, consequently, *for every solution*, $x_1(t), \cdots, x_n(t)$, *of* (3.1) *the characteristic number* (3.5) *will be finite. If*, on the other hand, *the integral* (3.4) *diverges*, one may assert that for every solution $x_1(t), \cdots, x_n(t)$ of (3.1) *the characteristic number*

$$C\left(x_1(t), \cdots, x_n(t) \mid \exp\left\{\int_{t_0}^{t} p(t)dt\right\}\right)$$

will be finite.

We shall give greater precision to the above results by examining more closely the lines of approach used in (L), p. 229 ff.

We shall need the following lemma, which will be stated without proof.

LEMMA 3.2. *Let $\psi(t)$ be an admissible function. Then*

$$C(x_1^2 + \cdots + x_n^2 \mid \psi(t)) = 2C(x_1, \cdots, x_n \mid \psi(t)), \tag{3.6}$$

provided at least one of the above characteristic numbers is finite.

In this connection we note that, given any set of functions $x_j(t)$ $(j = 1, \cdots, n)$, continuous and not all identically zero, there exist admissible functions $\psi(t)$ for which $C(x_1^2 + \cdots + x_n^2 \mid \psi(t))$ is finite. This follows from Theorem 2.1.

(8) (L), p. 229.

In view of Lemma 3.1 there is no loss of generality in considering the system (3.1) with coefficients subject to inequalities

$$|p_{i,j}(t)| \leqq b < \infty. \tag{3.7}$$

By Liapounoff's result ((L), p. 229 ff.), we merely know that every solution $x_j = x_j(t)$, $(j = 1, \cdots, n)$, of (3.1) has a finite characteristic number with respect to $\exp t$, provided (3.7) holds. We shall obtain bounds, depending on b, for the characteristic numbers of solutions.

Let $x_1, \cdots, x_n$ be a solution of (3.1). With Liapounoff, we employ the transformation

$$z_i = x_i e^{\lambda t} \qquad (\lambda \text{ an indeterminate number}), \tag{3.8}$$

which yields

$$\frac{dz_i}{dt} = p_{1,i}(t)z_1 + \cdots + (p_{i,i}(t) + \lambda)z_i + \cdots + p_{n,i}(t)z_n \qquad (i = 1, \cdots, n)$$

and, on writing $u = z_1^2 + \cdots + z_n^2$, we obtain

$$\frac{1}{2}\frac{du}{dt} = A(\lambda) + B(\lambda) = \Gamma(\lambda), \tag{3.9}$$

where

$$A(\lambda) = \sum_i (p_{i,i}(t) + \lambda)z_i^2, \quad B(\lambda) = \sum_{s \neq \sigma} (p_{s,\sigma}(t) + p_{\sigma,s}(t))z_s z_\sigma. \tag{3.9a}$$

Using the inequality of Schwartz, it can be proved that

$$\sum_{s \neq \sigma} |\alpha_s \alpha_\sigma| \leqq (n-1)(\alpha_1^2 + \cdots + \alpha_n^2) \qquad (s, \sigma = 1, \cdots, n)$$

for any numbers $\alpha_1, \cdots, \alpha_n$. Hence by (3.9a) and (3.7)

$$|B(\lambda)| \leqq 2b \sum_{s \neq \sigma} |z_s z_\sigma| \leqq 2(n-1)bu. \tag{3.10}$$

Define λ_0 as

$$\lambda_0 = (2n-1)b + \eta/2 \qquad (\eta > 0). \tag{3.11}$$

In consequence of (3.9a) and (3.7)

$$p_{i,i}(t) + \lambda \geqq \lambda_0 - b > 0 \qquad (\text{for } \lambda \geqq \lambda_0)$$

and

$$A(\lambda) \geqq (\lambda_0 - b)u > 0 \qquad (\text{for } \lambda \geqq \lambda_0).$$

By virtue of (3.10) and of the above, from (3.9) we deduce

$$\Gamma(\lambda) \geqq [(\lambda_0 - b) - 2(n-1)b]u = \frac{\eta}{2}u$$

for all $\lambda \geqq \lambda_0$ and for all $t \geqq t_0$. Whence by (3.9)

(3.12) $$u \geqq ce^{\eta t} \qquad (\lambda \geqq \lambda_0;\ \text{constant } c > 0).$$

It is similarly established that

$$\Gamma(\lambda) \leqq -\frac{\eta}{2}u \qquad (\text{for all } \lambda \leqq -\lambda_0),$$

which yields

(3.12a) $$u \leqq c_1 e^{-\eta t} \qquad (\lambda \leqq -\lambda_0;\ \text{constant } c_1 > 0).$$

By (3.8)

$$u = e^{2\lambda t}v, \qquad v = x_1^2 + \cdots + x_n^2;$$

here the x_j $(j=1, \cdots, n)$ are independent of λ and constitute a solution of (3.1). Thus, in view of (3.12) and (3.12a),

$$\lim e^{2\lambda t}v = +\infty \quad (\text{all } \lambda \geqq \lambda_0), \qquad \lim e^{2\lambda t}v = 0 \quad (\text{all } \lambda \leqq -\lambda_0),$$

where λ_0 is from (3.11). Inasmuch as η (>0) may be taken arbitrarily small, the above relations can be stated in the form

(3.13) $$\lim e^{2\lambda t}v = +\infty \ (\text{all } \lambda > (2n-1)b), \quad \lim e^{2\lambda t}v = 0 \ (\text{all } \lambda < -(2n-1)b).$$

An inequality

(3.14) $$\mu = C(v \mid e^t) > 2(2n-1)b$$

is impossible. In fact, suppose (3.14) holds. It is noted that

$$e^{(\mu-\epsilon)t}v = \left| e^{2[(2n-1)b+\epsilon]t}v \right| e^{[\mu-3\epsilon-2(2n-1)b]t}.$$

Choose ϵ subject to

$$0 < 3\epsilon \leqq \mu - 2(2n-1)b;$$

then $\mu - 3\epsilon - 2(2n-1)b \geqq 0$ and, in consequence of the first relation (3.13), we have

$$\lim e^{(\mu-\epsilon)t}v = \infty.$$

This is contrary to the relation $\mu = C(v \mid \exp t)$. Hence (3.14) is impossible and one has

$$\mu \leqq 2(2n-1)b.$$

It may be similarly shown that $\mu \geqq -2(2n-1)b$. Thus

$$(3.15) \qquad -2(2n-1)b \leqq C(x_1^2+\cdots+x_n^2 \mid e^t) \leqq 2(2n-1)b,$$

and, by virtue of Lemma 3.2,

$$(3.15a) \qquad -(2n-1)b \leqq C(x_1, \cdots, x_n \mid e^t) \leqq (2n-1)b.$$

We have obtained the following result.

THEOREM 3.1. *If in the differential system* (3.1) *the coefficients are continuous and*

$$|p_{i,j}(t)| \leqq b < \infty \qquad (i, j = 1, \cdots, n;\ t \geqq t_0),$$

the following may be asserted for every solution, x_j $(j=1, \cdots, n)$, *of* (3.1). *On writing* $v=x_1^2+\cdots+x_n^2$, *we have* (3.15), (3.15a) *and, more generally,* (3.13).

Consider now the case when $\lim p_{i,j}(t)=0$. *Let* $m(t)$ *be the "least" monotone non-increasing continuous function such that*

$$|p_{i,j}(t)| \leqq m(t) \qquad (i, j = 1, \cdots, n).$$

Existence of such a function is inferred without difficulty. Necessarily $\lim m(t)=0$. Moreover, inasmuch as we exclude the case, trivial from our point of view, when the $p_{i,j}(t)$ are identically zero, we have

$$(3.16) \qquad m(t) > 0.$$

Applying Theorem 3.1 (in particular (3.13)) to the interval (τ, ∞), where $\tau \geqq t_0$, it is concluded that

$$(3.17) \qquad \lim_{t\to\infty} e^{(\sigma(\tau)+\epsilon)t}v = \infty, \qquad \lim_{t\to\infty} e^{-(\sigma(\tau)+\epsilon)t}v = 0 \qquad (\text{all } \epsilon > 0),$$

with

$$(3.17a) \qquad \sigma(\tau) = 2(2n-1)m(\tau).$$

In fact, it is noted that for the interval $(\tau \leqq t < \infty)$ one may take $b=m(\tau)$.

In view of (3.16) there exists a number ϵ_0 (>0) so that the equation

$$\sigma(\tau) = \epsilon \qquad (0 < \epsilon \leqq \epsilon_0)$$

has a solution $\tau(\epsilon)$; necessarily

$$\sigma(\tau) \leqq \epsilon \qquad (\text{for } \tau \geqq \tau(\epsilon)),$$

inasmuch as $\sigma(\tau)$ of (3.17a) is monotone non-increasing just as $m(\tau)$ is. Thus, given ϵ such that $0<\epsilon\leqq\epsilon_0$, in (3.17) one may put $\tau=\tau(\epsilon)$, which will yield

$$\lim_{t\to\infty} e^{2\epsilon t}v = \infty, \qquad \lim_{t\to\infty} e^{-2\epsilon t}v = 0.$$

Whence

$$(3.18) \qquad C_\epsilon(v \mid e^t) = 0.$$

We have obtained

THEOREM 3.2. *Let the coefficients of the differential system* (3.1) *be continuous and*

$$\lim_{t\to\infty} p_{i,j}(t) = 0 \qquad (i, j = 1, \cdots, n).$$

For every solution $x_1, \cdots, x_n$ *of* (3.1) *we shall have* (3.18), *where* $v = x_1^2 + \cdots + x_n^2$.

This Theorem is applicable to the general system (3.1) envisaged at the beginning of this section. In fact, there are two cases at hand:

Case 3.19. The integral

$$\int_{t_0}^{\infty} p(t)dt \qquad (p(t) \text{ from } (3.2))$$

diverges.

Case 3.20. The above integral converges. That is, $p(t)\to 0$ sufficiently fast, as $t\to+\infty$.

Consider the Case 3.19 and let $\rho(t)$ be a continuous function such that

$$\rho(t) \geqq p(t). \tag{3.21}$$

Then the integral

$$\int_{t_0}^{\infty} \rho(t)dt$$

will diverge. Application of the transformation

$$\tau = \int_{t_0}^{t} \rho(t)dt \tag{3.22}$$

to (3.1) will yield

$$\frac{dx_i}{dt} = q_{1,i}(\tau)x_1 + \cdots + q_{n,i}(\tau)x_n \qquad (i = 1, \cdots, n), \tag{3.23}$$

with

$$q_{i,j}(\tau) = \frac{p_{i,j}(t)}{\rho(t)}.$$

The $q_{i,j}(\tau)$ will be continuous for $\tau \geqq 0$ and

$$|q_{i,j}(\tau)| \leqq \frac{p(t)}{\rho(t)} = r(\tau) \leqq 1 \qquad (i, j = 1, \cdots, n; \tau \geqq 0). \tag{3.23a}$$

If $\rho(t)$ is so chosen that

$$\lim_{t\to+\infty} \frac{p(t)}{\rho(t)} = 0, \tag{3.24}$$

one will have

$$\lim_{\tau\to\infty} r(\tau) = 0$$

and, in view of (3.23a), Theorem 3.2 will be applicable to the system (3.23), yielding the relation

$$C_s(x_1^2 + \cdots + x_n^2 \mid e^{\tau}) = 0 \tag{3.24a}$$

for every solution $x_1, \cdots, x_n$ of (3.23). By virtue of (3.22) this will imply that *in the Case* 3.19 *for every solution* $x_1, \cdots, x_n$ *of the differential system* (3.1) *we have*

$$C_s\left(x_1^2 + \cdots + x_n^2 \mid \exp\left\{\int_{t_0}^{t} \rho(t)dt\right\}\right) = 0, \tag{3.25}$$

provided $\rho(t)$ *satisfies* (3.21), (3.24).

The degree of precision afforded by an equality (3.25) *varies according to the choice of* $\rho(t)$ (subject to (3.21), (3.24))(*). If $\rho_1(t)$ and $\rho_2(t)$ are two functions, subject to (3.21), (3.24) and satisfying $\rho_1(t) \leqq \rho_2(t)$, of the two statements

$$C_s\left(v \mid \exp\left\{\int_{t_0}^{t} \rho_1(t)dt\right\}\right) = 0, \qquad C_s\left(v \mid \exp\left\{\int_{t_0}^{t} \rho_2(t)dt\right\}\right) = 0$$

the first one will be as precise as or more precise than the second. Clearly there exists no particular function $\rho(t)$, satisfying (3.21) and (3.24), for which (3.25) represents the most precise result.

Consider now Case 3.20. Theorem 3.2 will be applicable directly to the system (3.1). Let $\rho(t)$ be positive and continuous (for $t \geqq t_0$) and let $\rho(t) \to 0$ (as $t \to \infty$) sufficiently slowly so that

$$\lim \frac{p(t)}{\rho(t)} = 0, \qquad \int_{t_0}^{\infty} \rho(t)dt \text{ diverges.} \tag{3.26}$$

Applying to the system (3.1) the transformation (3.22), we repeat the steps leading from (3.22) to (3.24a). Thus, *in the Case* 3.20 (and, more generally, when $\lim p_{i,j}(t) = 0$) *for every solution* $x_1, \cdots, x_n$ *of the differential system* (3.1) *we have*

$$C_s\left(x_1^2 + \cdots + x_n^2 \mid \exp\left\{\int_{t_0}^{t} \rho(t)dt\right\}\right) = 0,$$

where $\rho(t)$ (>0) *is taken continuous* (for $t \geqq t_0$) *and subject to* (3.26). With respect to the degree of precision presented by the above result a statement may be made analogous to that for the Case 3.19.

(*) If one takes $\rho(t) > 0$, the condition (3.21) will be non-essential.

The following is a simple application of the above result. Suppose

$$|p_{i,j}(t)| \leqq at^{-\alpha} \qquad (a > 0, \alpha > 1; t \geqq t_0 > 0);$$

on taking $\rho(t) = t^{-1}$ we then obtain

$$C_s(x_1^2 + \cdots + x_n^2 \mid t/t_0) = 0$$

for every solution $x_1, \cdots, x_n$ of (3.1).

4. **Growth of solutions for linear systems.** In the previous section we obtained results relating to properties of "growth," at $t = +\infty$, of solutions of linear systems (3.1). We shall now apply a different method, based on "product integration," for the purpose of establishing some further results regarding properties of solutions of (3.1) for t large.

Using matrix notation the system (3.1) may be written as

$$X^{(1)}(t) = X(t)P(t), \qquad X(t) = (x_{i,j}(t)), \qquad P(t) = (p_{i,j}(t)), \tag{4.1}$$

$i, j = 1, \cdots, n$. Here and in the sequel the element $a_{i,j}$ displayed in a matrix $(a_{i,j})$ will be the one in the ith row and jth column. If $X(t)$ satisfies (4.1), the elements of any row in $X(t)$ will constitute a solution of (3.1). With $t > t_0$, we subdivide the interval (t_0, t) into N equal subintervals and call their end points

$$t_\nu = t_0 + \nu\Delta \quad (\nu = 0, 1, \cdots, N), \qquad t = t_N,$$

where

$$\Delta = \frac{t - t_0}{N}. \tag{4.2}$$

We then form the matrix

$$X_N = (x_{N:i,j}) = (I + \Delta P(t_0))(I + \Delta P(t_1)) \cdots (I + \Delta P(t_{N-1})), \tag{4.3}$$

where $I = (\delta_{i,j})$ is the identity matrix. In consequence of the known theory of "product integration" the limit

$$\lim_{N \to \infty} X_N = X(t) = (x_{i,j}(t)) \tag{4.4}$$

exists for all $t \geqq t_0$ and the matrix $X(t)$ will satisfy (4.1); moreover, $X(t_0) = I$. We have taken equal subintervals for convenience.

By (4.3)

$$X_N = I + \Delta \sum_{\nu_1=0}^{N-1} P(t_{\nu_1}) + \Delta^2 \sum_{\nu_1 < \nu_2} P(t_{\nu_1})P(t_{\nu_2}) + \cdots$$
$$+ \Delta^m \sum_{\nu_1 < \cdots < \nu_m} P(t_{\nu_1})P(t_{\nu_2}) \cdots P(t_{\nu_m}) + \cdots + \Delta^N P(t_0)P(t_1) \cdots P(t_{N-1});$$

here $\nu_j = 0, 1, \cdots, N-1$. Thus

$$X_N - I = \sum_{m=1}^{N} \Delta^m L_m, \qquad L_m = (l_{m:i,j}), \tag{4.5}$$

where the element $l_{m:i,j}$ in the matrix L_m is of the form

$$\begin{aligned} l_{m:i,j} &= \text{element in the } i\text{th row and } j\text{th} \\ &\quad \text{column of } \textstyle\sum_{\nu_1<\cdots<\nu_m} P(t_{\nu_1}) \cdots P(t_{\nu_m}) \\ &= \sum_{\nu_1<\cdots<\nu_m} \sum_{\tau_1,\cdots,\tau_{m-1}=1}^{n} p_{i,\tau_1}(t_{\nu_1}) p_{\tau_1,\tau_2}(t_{\nu_2}) \cdots p_{\tau_{m-1},j}(t_{\nu_m}). \end{aligned} \tag{4.5a}$$

Let $p(t)$ be a continuous function such that

$$|\, p_{i,j}(t)\,| \leqq p(t) \qquad (i, j = 1, \cdots, n). \tag{4.6}$$

In consequence of (4.5a) it is concluded that

$$\begin{aligned} |\, \Delta^m l_{m:i,j}\,| &\leqq \sum_{\nu_1<\cdots<\nu_m} \sum_{\tau_1,\cdots,\tau_{m-1}=1}^{n} (\Delta p(t_{\nu_1}))(\Delta p(t_{\nu_2})) \cdots (\Delta p(t_{\nu_m})) \\ &= \sum_{\nu_1,\cdots} (\Delta p(t_{\nu_1})) \cdots (\Delta p(t_{\nu_m})) \sum_{\tau_1,\cdots} 1 \\ &= \sum_{\nu_1<\cdots<\nu_m} (\Delta p(t_{\nu_1})) \cdots (\Delta p(t_{\nu_m})) n^{m-1} \\ &= \frac{1}{n} \sum_{\nu_1<\cdots<\nu_m} (n\Delta p(t_{\nu_1}))(n\Delta p(t_{\nu_2})) \cdots (n\Delta p(t_{\nu_m})). \end{aligned} \tag{4.7}$$

By (4.5) and (4.7)

$$\begin{aligned} n\,|\, x_{N:i,j} - \delta_{i,j}\,| &\leqq n \sum_{m=1}^{N} |\, \Delta^m l_{m:i,j}\,| \\ &\leqq \sum_{m=1}^{N} \sum_{\nu_1<\cdots<\nu_m} (n\Delta p(t_{\nu_1}))(n\Delta p(t_{\nu_2})) \cdots (n\Delta p(t_{\nu_m})), \end{aligned} \tag{4.8}$$

where $\nu_j = 0, \cdots, N-1$. With the aid of the identity

$$-1 + (1 + c_0)(1 + c_1) \cdots (1 + c_{N-1}) = \sum_{m=1}^{N} \sum_{\nu_1<\cdots<\nu_m} c_{\nu_1} c_{\nu_2} \cdots c_{\nu_m}$$

$(\nu_j = 0, \cdots, N-1)$ from (4.8) we infer

$$n\,|\, x_{N:i,j} - \delta_{i,j}\,| \leqq -1 + \prod_{\tau=0}^{N-1} (1 + n\Delta p(t_n)).$$

Now, inasmuch as $n\Delta p(t_n) \geqq 0$ and since

$$1 + u \leqq e^u \qquad (\text{for } u \geqq 0),$$

we obtain

$$n\,|\,x_{N:i,j} - \delta_{i,j}\,| \leqq -1 + \exp n \sum_{\tau=0}^{N-1} \Delta p(t_\tau). \tag{4.8a}$$

From this in the limit, as $N \to \infty$, it is deduced that

$$|\,x_{i,j}(t) - \delta_{i,j}\,| \leqq \frac{1}{n}\left[-1 + \exp\left\{n \int_{t_0}^{t} p(t)dt\right\}\right] = \alpha(t); \tag{4.9}$$

here the $x_{i,j}(t)$ $(i, j = 1, \cdots, n)$ are the elements of the matrix (4.4). The best result is presented by (4.9) when $p(t)$ of (4.6) is taken "least."

Let

$$t_0 \leqq t \leqq \tau. \tag{4.10}$$

We subdivide the interval (t, τ) into N equal subintervals, with end points given by

$$t = \tau_N < \tau_{N-1} < \cdots < \tau_1 < \tau_0 = \tau,$$

where

$$\tau_\nu = \tau + \nu\delta, \qquad \delta = \frac{t - \tau}{N} = -\,|\,\delta\,|. \tag{4.10a}$$

If the matrix

$$Z_N = (z_{N:i,j}) = (I + \delta P(\tau_0))(I + \delta P(\tau_1)) \cdots (I + \delta P(\tau_{N-1}))$$

is formed, in consequence of the theory of product integration it is deduced that the limit

$$\lim_N Z_N = Z(t) = (z_{i,j}(t)) \qquad (t_0 \leqq t \leqq \tau) \tag{4.11}$$

exists and satisfies (4.1); we have

$$Z(\tau) = I. \tag{4.11a}$$

In place of (4.8) we now have inequalities obtained by replacing

$$x_{N:i,j},\ t_\nu,\ \Delta$$

by

$$z_{N:i,j},\ \tau_\nu,\ -\delta,$$

respectively. We had $\Delta > 0$, but now $\delta < 0$. Accordingly, in place of (4.8a) it is established that

$$n\,|\,z_{N:i,j} - \delta_{i,j}\,| \leqq -1 + \exp\left\{-n \sum_{\tau=0}^{N-1} \delta p(\tau_\tau)\right\}.$$

In the limit, as $N \to \infty$, this will yield the following inequalities for the functions $z_{i,j}(t)$ of (4.11a)

$$n\,|\,z_{i,j}(t) - \delta_{i,j}\,| \leqq -1 + \exp\left\{-n\int_{\tau}^{t} p(t)dt\right\} \qquad (t_0 \leqq t \leqq \tau). \tag{4.12}$$

When the integral

$$\int_{t_0}^{\infty} p(t)dt \qquad \text{(cf. (4.6))} \tag{4.13}$$

converges, the following is observed. The matrix

$$X(t) = (x_{i,j}(t)) = \lim_{\tau \to +\infty} Z(t)$$

satisfies the system (4.1); moreover,

$$X(+\infty) = I,$$

and, by (4.12),

$$|\,x_{i,j}(t) - \delta_{i,j}\,| \leqq \frac{1}{n}\left[-1 + \exp\left\{n\int_{t}^{\infty} p(t)dt\right\}\right] = \beta(t) \qquad (t_0 \leqq t < \infty). \tag{4.14}$$

Accordingly one may formulate the theorem.

THEOREM 4.1. *The following may be stated for the solutions of the differential system* (3.1), *whose equivalent matrix form is* (4.1).

Let $p(t)$ be a continuous function such that

$$|\,p_{i,j}(t)\,| \leqq p(t).$$

The matrix solution $X(t) = (x_{i,j}(t))$, for which $X(t_0) = I$, satisfies the inequalities (4.9). *In the case when the integral*

$$\int_{t_0}^{\infty} p(t)dt$$

converges, there exists a matrix solution $X(t)$, for which $X(+\infty) = I$; the elements of $X(t)$ will satisfy the inequalities (4.14).

It is noted that, when (4.13) converges, the function $\alpha(t)$ of (4.9) is uniformly bounded; in fact

$$\alpha(t) \leqq \alpha(\infty) < \infty \qquad (t_0 \leqq t < \infty).$$

When (4.13) diverges $\alpha(t)$ is monotone increasing and

$$\lim_{t \to \infty} \alpha(t) = +\infty;$$

moreover, in this case, the faster

$$\phi(t) = \int_{t_0}^{t} p(t)dt$$

approaches $+\infty$, as $t \to \infty$, the faster will $\alpha(t)$ approach $+\infty$ with t.

When (4.13) converges, the function $\beta(t)$ of (4.14) will approach zero, as $t \to +\infty$. In this case, the faster $p(t) \to 0$ (as $t \to \infty$), the faster will $\beta(t)$ approach zero with $1/t$.

Comparison of the results of Theorems 4.1, 3.1 leads to the conclusion that the two theorems overlap and, in a sense, supplement each other.

COROLLARY 4.1. *With $p(t)$ from* (4.6) *and*

$$\int_{t_0}^{\infty} p(t)dt$$

convergent, every solution $Y(t) = (y_{i,j}(t))$ (defined for $t \geqq t_0$) of (4.1) *has the properties*:

$$\lim Y(t) = C = (c_{i,j}) \qquad (as\ t \to \infty), \tag{4.15}$$

where the $c_{i,j}$ are constants;

$$| y_{i,j}(t) - c_{i,j} | \leqq nc\beta(t),$$

with $\beta(t)$ from (4.14) *and the constant c from the inequalities*

$$| c_{i,j} | \leqq c \qquad (i, j = 1, \cdots, n). \tag{4.15a}$$

In fact,

$$Y(t) = CX(t), \qquad C = (c_{i,j}),$$

where the $c_{i,j}$ are some constants and $X(t)$ is a matrix solution of (4.1) for which $X(+\infty) = I$. The solution $X(t)$, with the stated property, exists in consequence of Theorem 4.1. The relation (4.15) follows at once. Now

$$Y(t) - C = (y_{i,j}(t) - c_{i,j}) = C(X(t) - I)$$

and

$$y_{i,j}(t) - c_{i,j} = \sum_{r=1}^{n} c_{i,r}(x_{r,j}(t) - \delta_{r,j}).$$

Thus, by (4.15a) and (4.14)

$$| y_{i,j}(t) - c_{i,j} | \leqq nc\beta(t),$$

which establishes the corollary.

Let us consider the case when there exists a continuous function $\rho(t)$ such that

$$p_{i,j}(t) \geqq \rho(t) \geqq 0 \qquad (i, j = 1, \cdots, n). \tag{4.16}$$

We repeat the developments leading up to (4.5a). By (4.5a) and (4.16)

$$\Delta l_{1:i,j} = \sum_{r_1=0}^{N-1} \Delta p_{i,j}(t_{r_1}) \geqq \sum_{r_1=0}^{N-1} \Delta p(t_{r_1}), \qquad \Delta^m l_{m:i,j} \geqq 0.$$

Whence, in view of (4.5),

$$x_{N:i,j} - \delta_{i,j} = \Delta l_{1:i,j} + \sum_{m=2}^{N} \Delta^m l_{m:i,j} \geqq \Delta l_{1:i,j} \geqq \sum_{r_1=0}^{N-1} \Delta\rho(t_{r_1})$$

and, in the limit (as $N \to \infty$),

$$x_{i,j}(t) - \delta_{i,j} \geqq \lim_{N} \sum_{r_1=0}^{N-1} \Delta p(t_{r_1}) = \int_{t_0}^{t} \rho(t)dt. \tag{4.17}$$

Hence we may state the theorem.

THEOREM 4.2. *Suppose* (4.16) *holds. Then the* $x_{i,j}(t)$ *of the matrix solution* $X(t) = (x_{i,j}(t))$ *of* (4.1), *with* $X(t_0) = I$, *will satisfy the inequalities* (4.17).

COROLLARY 4.2. *Suppose that the coefficients in* (4.1) *satisfy*

$$p_{i,j}(t) \leqq -\rho(t) < 0 \qquad (i, j = 1, \cdots, n),$$

where $\rho(t)$ *is continuous. With* $X(t) = (x_{i,j}(t))$ *denoting a solution of* (4.1), *reducing to* I *for* t_0, *the elements* $\bar{x}_{i,j}(t)$ *of the inverse of* $X(t)$ *will satisfy the inequalities*

$$\bar{x}_{i,j}(t) \geqq \delta_{i,j} + \int_{t_0}^{t} \rho(t)dt \qquad (i, j = 1, \cdots, n).$$

The corollary is established on noting that

$$\bar{X}(t) = (\bar{x}_{i,j}(t))$$

satisfies

$$\bar{X}^{(1)}(t) = -P(t)\bar{X}(t),$$

whenever $X(t)$ is a matrix solution of (4.1).

Further results will be obtained with the aid of product integration and of the identity

$$\begin{aligned} &h_{i,r_1}(t_{r_1}) + h_{r_1,r_2}(t_{r_2}) + \cdots + h_{r_{m-2},r_{m-1}}(t_{r_{m-1}}) + h_{r_{m-1},j}(t_{r_m}) \\ &= h_{i,j}(t) + [h_{j,i}(t) - h_{j,i}(t_{r_m})] + [h_{r_{m-1},i}(t_{r_m}) - h_{r_{m-1},i}(t_{r_{m-1}})] + \cdots \\ &\quad + [h_{r_2,i}(t_{r_2}) - h_{r_2,i}(t_{r_2})] + [h_{r_1,i}(t_{r_2}) - h_{r_1,i}(t_{r_1})], \end{aligned} \tag{4.18}$$

where we have put

$$h_{i,j}(t) = h_i(t) - h_j(t) \qquad (i, j = 1, \cdots, n). \tag{4.18a}$$

Suppose the $h_i(t)$ are continuous for $t \geqq t_0$ and the $h_i^{(1)}(t)$ $(i=1, \cdots, n)$ exist. If we have

$$h_1^{(1)}(t) \geqq h_2^{(1)}(t) \geqq \cdots \geqq h_n^{(1)}(t), \tag{4.19}$$

then

$$h_{\tau,i}^{(1)}(t) = h_\tau^{(1)}(t) - h_i^{(1)}(t) \leqq 0 \qquad (i = 1, \cdots, \mu; \tau = 1, \cdots, n), \tag{4.19a}$$

where μ is from the inequalities

$$h_1^{(1)}(t) = h_2^{(1)}(t) = \cdots = h_\mu^{(1)}(t) \geqq h_{\mu+1}^{(1)}(t) \geqq \cdots . \tag{4.19b}$$

Thus, under (4.19), $1 \leqq \mu \leqq n$. With $t_0 \leqq t' \leqq t''$, from (4.19a) one would obtain

$$h_{\tau,i}(t'') - h_{\tau,i}(t') = \int_{t'}^{t''} h_{\tau,i}^{(1)}(t)dt \leqq 0 \quad (i = 1, \cdots, \mu; \tau = 1, \cdots, n). \tag{4.19c}$$

Whence, *if* (4.19) (*see* (4.19b)) *holds, the identity* (4.18) *will yield*

$$h_{i,\tau_1}(t_{\tau_1}) + h_{\tau_1,\tau_2}(t_{\tau_2}) + \cdots + h_{\tau_{m-2},\tau_{m-1}}(t_{\tau_{m-1}}) + h_{\tau_{m-1},j}(t_{\tau_m}) \leqq h_{i,j}(t) \qquad (i = 1, \cdots, \mu; j = 1, \cdots, n), \tag{4.20}$$

whenever

$$t_0 \leqq t_{\tau_1} \leqq t_{\tau_2} \leqq \cdots \leqq t_{\tau_m} \leqq t. \tag{4.20a}$$

If instead of (4.19) we have

$$h_1^{(1)}(t) \leqq h_2^{(1)}(t) \leqq \cdots \leqq h_n^{(1)}(t), \tag{4.21}$$

then for some number η, such that $1 \leqq \eta \leqq n$, the inequalities

$$h_1^{(1)}(t) \leqq \cdots \leqq h_\eta^{(1)}(t) = h_{\eta+1}^{(1)}(t) = \cdots = h_n^{(1)}(t) \tag{4.21a}$$

will hold. It will follow that

$$h_{\tau,i}^{(1)}(t) \leqq 0 \qquad (i = \eta, \eta+1, \cdots, n; \tau = 1, \cdots, n).$$

This, in turn, would imply

$$h_{\tau,i}(t'') - h_{\tau,i}(t') \leqq 0 \qquad (i = \eta, \eta+1, \cdots, n; \tau = 1, \cdots, n)$$

for any values t', t'' such that $t_0 \leqq t' \leqq t''$. Hence, on taking note of (4.18), it is observed that, *if* (4.21) (*see* (4.21a)) *holds, one has*

$$h_{i,\tau_1}(t_{\tau_1}) + h_{\tau_1,\tau_2}(t_{\tau_2}) + \cdots + h_{\tau_{m-2},\tau_{m-1}}(t_{\tau_{m-1}}) + h_{\tau_{m-1},j}(t_{\tau_m}) \leqq h_{i,j}(t) \qquad (i = \eta, \eta+1, \cdots, n; j = 1, \cdots, n), \tag{4.22}$$

whenever (4.20a) *is satisfied.*

Let $w(t)$, $h_j(t)$ $(j=1, \cdots, n)$ be functions such that (4.19) holds and

$$|p_{i,j}(t)| \leqq w(t)e^{h_i(t)-h_j(t)} \qquad (i, j = 1, \cdots, n), \tag{4.23}$$

the function $w(t)$ ($\geqq 0$) being continuous for $t \geqq t_0$. Such functions exist in all cases. Repeating the developments leading to (4.5a) we now obtain

$$\begin{aligned} |\Delta^m l_{m:i,j}| \leqq \sum_{\nu_1<\cdots<\nu_m} \sum_{\tau_1,\cdots,\tau_{m-1}=1}^{n} (\Delta w(t_{\nu_1}))(\Delta w(t_{\nu_2})) \cdots (\Delta w(t_{\nu_m})) \\ \cdot \exp\left[h_{i,\tau_1}(t_{\nu_1}) + h_{\tau_1,\tau_2}(t_{\nu_2}) + \cdots + h_{\tau_{m-2},\tau_{m-1}}(t_{\nu_{m-1}}) + h_{\tau_{m-1},j}(t_{\nu_m})\right]. \end{aligned} \tag{4.24}$$

Inasmuch as (4.20a) is satisfied, application of (4.20) to (4.24) is possible, yielding

$$|\Delta^m l_{m:i,j}| \leqq e^{h_{i,j}(t)} \sum_{\nu_1<\cdots<\nu_m} \sum_{\tau_1,\cdots} (\Delta w(t_{\nu_1})) \cdots (\Delta w(t_{\nu_m})) \tag{4.24a}$$

for $i = 1, \cdots, \mu$ (μ from (4.19b)). Using the steps employed in (4.7), from (4.24a) it is deduced that

$$|\Delta^m l_{m:i,j}| \leqq e^{h_{i,j}(t)} \frac{1}{n} \sum_{\nu_1<\cdots<\nu_m} (n\Delta w(t_{\nu_1})) \cdots (n\Delta w(t_{\nu_m}))$$

($i = 1, \cdots, \mu$; $j = 1, \cdots, n$; $0 \leqq \nu_\nu \leqq N-1$). By (4.5) and the above (see (4.8)) one has

$$n|x_{N:i,j} - \delta_{i,j}|e^{-h_{i,j}(t)} \leqq \sum_{m=1}^{N} \sum_{\nu_1,\cdots} (n\Delta w(t_{\nu_1})) \cdots (n\Delta w(t_{\nu_m}))$$

($i = 1, \cdots, \mu$) and, with the aid of the identity subsequent to (4.8), we find

$$n|x_{N:i,j} - \delta_{i,j}|e^{-h_{i,j}(t)} \leqq -1 + \prod_{\tau=0}^{N-1} (1 + n\Delta w(t_\tau)) \qquad (i = 1, \cdots, \mu).$$

Thus

$$n|x_{N:i,j} - \delta_{i,j}|e^{-h_{i,j}(t)} \leqq -1 + \exp n \sum_{\tau=0}^{N-1} \Delta w(t_\tau) \qquad (i = 1, \cdots, \mu),$$

and, in the limit (as $N \to \infty$),

$$|x_{i,j}(t) - \delta_{i,j}| \leqq \frac{1}{n} e^{h_i(t)-h_j(t)}\left[-1 + \exp\left\{n\int_{t_0}^{t} w(t)dt\right\}\right] \tag{4.25}$$
$$(i = 1, \cdots, \mu; j = 1, \cdots, n).$$

Let continuous $w(t)$ ($\geqq 0$) and the $h_j(t)$ be such that (4.21) holds and the inequalities (4.23) are satisfied. On making use of (4.22), we now obtain

$$|x_{i,j}(t) - \delta_{i,j}| \leqq \frac{1}{n} e^{h_i(t)-h_j(t)}\left[-1 + \exp\left\{n\int_{t_0}^{t} w(t)dt\right\}\right] \tag{4.26}$$
$$(i = \eta, \eta+1, \cdots, n; j = 1, \cdots, n),$$

where η is from (4.21a). As in (4.25) the $x_{i,j}(t)$ are the elements of the matrix $X(t)$, satisfying the system (4.1); $X(t_0)=I$.

Inequalities (4.25) refer to the first μ rows of $X(t)$, while (4.26) refers to the last $n-\eta+1$ rows. We recall that the elements in any row of $X(t)$ constitute a solution of the system (3.1). The $h_j(t)$ in (4.25) and (4.26) do not need to be the same.

Inequalities of the form (4.23) imply, in particular, that

$$w(t) \geqq |\, p_{i,i}(t)\,| \qquad (i = 1, \cdots, n).$$

Thus, it is not always possible to choose $w(t)$ so that the integral

$$\int_{t_0}^{\infty} w(t)dt \tag{4.27}$$

converges. However, for some systems (4.1) one may choose $w(t)$ so that (4.27) converges. Suppose now that the latter is the case. Let $w(t)$ be a continuous function such that (4.27) converges, while

$$|\, p_{i,j}(t)\,| \leqq w(t)e^{h_i(t)-h_j(t)} \qquad (i, j = 1, \cdots, n), \tag{4.28}$$

where the $h_j(t)$ are such that either (4.19) or (4.21) holds. We recall the notation introduced in connection with (4.10)–(4.10a), as well as the subsequent developments leading to (4.14). The formula corresponding to (4.8) will be

$$n\,|\, z_{N:i,j} - \delta_{i,j}\,| \leqq n\sum_{m=1}^{N}(-\delta)^m\,|\, l_{m:i,j}\,|. \tag{4.29}$$

On the other hand, making the replacements indicated subsequent to (4.11a), in view of (4.5a) it is concluded that

$$(-\delta)^m\,|\, l_{m:i,j}\,| \leqq \sum_{\nu_1<\cdots<\nu_m}\ \sum_{\lambda_1,\cdots,\lambda_{m-1}=1}^{n}(-\delta p_{i,\lambda_1}(\tau_{\nu_1}))(-\delta p_{\lambda_1,\lambda_2}(\tau_{\nu_2}))\cdots(-\delta p_{\lambda_{m-1},j}(\tau_{\nu_m}))$$

$(0\leqq\nu_\sigma\leqq N-1)$. In consequence of (4.28) this will imply (cf. (4.24))

$$\begin{aligned}(-\delta)^m\,|\, l_{m:i,j}\,| &\leqq \sum_{\nu_1<\cdots<\nu_m}\ \sum_{\lambda_1,\cdots,\lambda_{m-1}=1}^{n}(-\delta w(\tau_{\nu_1}))(-\delta w(\tau_{\nu_2}))\cdots(-\delta w(\tau_{\nu_m}))\\ &\quad\cdot\exp\left[h_{i,\lambda_1}(\tau_{\nu_1}) + h_{\lambda_1,\lambda_2}(\tau_{\nu_2}) + \cdots + h_{\lambda_{m-2},\lambda_{m-1}}(\tau_{\nu_{m-1}}) + h_{\lambda_{m-1},j}(\tau_{\nu_m})\right]\end{aligned} \tag{4.30}$$

$(i, j=1, \cdots, n;\ t<\tau_{\nu_m}<\tau_{\nu_{m-1}}<\cdots<\tau_{\nu_1}\leqq\tau)$. Under (4.19), in place of (4.19c) we obtain

$$h_{\lambda,i}(t'') - h_{\lambda,i}(t') = \int_{t'}^{t''} h_{\lambda,i}^{(1)}(t)dt \leqq 0 \quad (i = \eta', \cdots, n;\ \lambda = 1, \cdots, n) \tag{4.31}$$

for $t \leqq t'' \leqq t' \leqq \tau$; here η' $(1 \leqq \eta' \leqq n)$ is, from the inequalities, valid when (4.19) holds—that is,

$$(4.31a) \quad h_1^{(1)}(t) \geqq h_2^{(1)}(t) \geqq \cdots \geqq h_{\eta'}^{(1)}(t) = h_{\eta'+1}^{(1)}(t) = \cdots = h_n^{(1)}(t).$$

On the other hand, under (4.21),

$$(4.32) \qquad h_{\lambda,i}(t'') - h_{\lambda,i}(t') \leqq 0 \qquad (i = 1, \cdots, \mu'; \lambda = 1, \cdots, n)$$

for $t \leqq t'' \leqq t' \leqq \tau$, the integer μ' being from the inequalities

$$(4.32a) \quad h_1^{(1)}(t) = h_2^{(1)}(t) = \cdots = h_{\mu'}^{(1)}(t) \leqq h_{\mu'+1}^{(1)}(t) \leqq \cdots \leqq h_n^{(1)}(t).$$

By virtue of (4.18)

$$\begin{aligned} W &= h_{i,\lambda_1}(\tau_{\nu_1}) + h_{\lambda_1,\lambda_2}(\tau_{\nu_2}) + \cdots + h_{\lambda_{m-1},j}(\tau_{\nu_m}) \\ &= h_{i,j}(t) + [h_{j,i}(t) - h_{j,i}(\tau_{\nu_m})] + [h_{\lambda_{m-1},i}(\tau_{\nu_m}) - h_{\lambda_{m-1},i}(\tau_{\nu_{m-1}})] + \cdots \\ &\quad + [h_{\lambda_1,i}(\tau_{\nu_2}) - h_{\lambda_1,i}(\tau_{\nu_1})]. \end{aligned} \tag{4.33}$$

Inasmuch as the τ_j satisfy the inequalities of (4.30), on taking account of the statements in connection with (4.31) and (4.32), the following is inferred. For $j = 1, \cdots, n$ and $t_0 \leqq t \leqq \tau$

$$(4.34) \qquad W \leqq h_{i,j}(t) \begin{cases} i = \eta', \cdots, n & \text{(under (4.19); cf. (4.31a)),} \\ i = 1, \cdots, \mu' & \text{(under (4.21); cf. (4.32a)).} \end{cases}$$

Application of (4.34) will yield from (4.30)

$$(-\delta)^m \left| l_{m:i,j} \right| \leqq e^{h_{i,j}(t)} \sum_{\nu_1 < \cdots < \nu_m} \sum_{\lambda_1, \cdots} (-\delta w(\tau_{\nu_1})) \cdots (-\delta w(\tau_{\nu_m}))$$

for values of i as indicated in (4.34) and for $t_0 \leqq t \leqq \tau$. Whence, by virtue of (4.29),

$$\begin{aligned} n \left| z_{N:i,j} - \delta_{i,j} \right| e^{-h_{i,j}(t)} &\leqq \sum_{m=1}^{N} \sum_{\nu_1 < \cdots} (-n\delta w(\tau_{\nu_1})) \cdots (-n\delta w(\tau_{\nu_m})) \\ &= -1 + \prod_{r=0}^{N-1} (1 - n\delta w(\tau_r)) \\ &\leqq -1 + \exp - n \sum_{r=0}^{N-1} \delta w(\tau_r) \end{aligned}$$

for $i = \eta', \cdots, n$, under (4.19), or for $i = 1, \cdots, \mu'$, under (4.21). Accordingly, in the limit (as $N \to \infty$) one has

$$\left| z_{i,j}(t) - \delta_{i,j} \right| \leqq \frac{1}{n} e^{h_i(t) - h_j(t)} \left[-1 + \exp \left\{ -n \int_\tau^t w(t)\,dt \right\} \right] \tag{4.35}$$

$$(i = \eta', \cdots, n \text{ or } i = 1, \cdots, \mu'; j = 1, \cdots, n; t_0 \leqq t \leqq \tau).$$

The matrix $(z_{i,j}(t))$ satisfies the system (4.1); moreover, $(z_{i,j}(\tau)) = (\delta_{i,j})$.

Existence of the limits, as $\tau \to \infty$, cannot be asserted for all the elements of the matrix $(z_{i,j}(t))$. However, the limits

$$\lim_{\tau} z_{i,j}(t) = x_{i,j}(t) \qquad (= \eta', \cdots, n; \text{ under (4.19)}), \tag{4.36}$$

$$\lim_{\tau} z_{i,j}(t) = x_{i,j}(t) \qquad (i = 1, \cdots, \mu'; \text{ under (4.21)}) \tag{4.36a}$$

will exist for $j = 1, \cdots, n$; moreover, in consequence of (4.35)

$$| x_{i,j}(t) - \delta_{i,j} | \leqq \frac{1}{n} e^{h_i(t) - h_j(t)} \left[-1 + \exp \left\{ n \int_t^{\infty} w(t)dt \right\} \right] \tag{4.37}$$

for values of i as indicated in (4.36), (4.36a), for $j = 1, \cdots, n$ and for $t_0 \leqq t < \infty$, it being supposed that (4.27) converges. For any i, for which the above is asserted, the set of n functions

$$x_{i,1}(t), x_{i,2}(t), \cdots, x_{i,n}(t) \tag{4.37a}$$

constitutes a solution of (3.1).

The above results may be summed as follows.

THEOREM 4.3. *Let $w(t)$ $(\geqq 0)$, $h_j(t)$ $(j = 1, \cdots, n)$ be functions continuous for $t \geqq t_0$, such that*

$$| p_{i,j}(t) | \leqq w(t) \exp [h_i(t) - h_j(t)] \qquad (i, j = 1, \cdots, n),$$

and such that either (4.19) or (4.21) is satisfied, the $h_j(t)$ being derivable.

If $X(t) = (x_{i,j}(t))$ is the matrix solution of (4.1), for which $X(t_0) = I$, then

$$| x_{i,j}(t) - \delta_{i,j} | \leqq \frac{1}{n} e^{h_i(t) - h_j(t)} \left[-1 + \exp \left\{ n \int_{t_0}^{t} w(t)dt \right\} \right]$$

$(j = 1, \cdots, n)$ for $i = 1, \cdots, \mu$, under (4.19) (cf. (4.19b)), and for $i = \eta, \cdots, n$, under (4.21) (cf. (4.21a)).

Suppose $w(t)$ can be so selected that the integral of $w(t)dt$ extended over (t_0, ∞) converges, while the $h_j(t)$ are functions as described above, for which either (4.19) or (4.21) holds. One may then assert that there exists a matrix solution $X(t) = x_{i,j}(t))$ of (4.1) with the following properties. If (4.19) holds, each of the last $n - \eta' + 1$ rows of $X(t)$ will constitute a solution of the system (3.1). When (4.21) holds, each of the first μ' rows of $X(t)$ will be a solution of (3.1). In this connection, η', μ' are from (4.31a) and (4.32a), respectively. The constituent elements $x_{i,j}(t)$ of the solutions of (3.1), just referred to, will satisfy inequalities (4.37) for $j = 1, \cdots, n$ and for $t_0 \leqq t < \infty$.

5. **Growth of solutions for nonlinear systems.** We consider now nonlinear differential systems of the form important in classical dynamics,

$$\frac{dx_i}{dt} = f_i(t; x_1, \cdots, x_n) = l_i + q_i \qquad (i = 1, \cdots, n), \tag{5.1}$$

where

$$l_i = p_{1,i}(t)x_1 + \cdots + p_{n,i}(t)x_n, \tag{5.1a}$$

the $p_{s,i}(t)$ being functions of the same type as before, while

$$q_i = \sum_{j_1+\cdots+j_n=2} p_i^{j_1,\cdots,j_n}(t; x)x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n} \quad (j_1, \cdots, j_n \geqq 0). \tag{5.1b}$$

The coefficients in (5.1b) may depend on $(x) = (x_1, \cdots, x_n)$:

$$p_i^{j_1,\cdots,j_n}(t; x) = p_i^{j_1,\cdots,j_n}(t) + r_i^{j_1,\cdots,j_n}(t; x), \tag{5.1c}$$

where the

$$p_i^{j_1,\cdots,j_n}(t) \qquad (j_1 + \cdots + j_n = 2; j_1, \cdots, j_n \geqq 0)$$

are independent of $x_1, \cdots, x_n$ and are continuous for $t \geqq t_0$. On the other hand, *the*

$$r_i^{j_1,\cdots,j_n}(t; x) = r_i^{j_1,\cdots,j_n}(t; x_1, \cdots, x_n)$$

are continuous in $x_1, \cdots, x_n$, *for every* $t \geqq t_0$, *when*

$$t_0 \leqq t < \infty, \qquad |x_j| \leqq H \qquad (j = 1, \cdots, n; H > 0); \tag{5.2}$$

moreover, it is assumed that

$$|r_i^{j_1,\cdots,j_n}(t; x)| \leqq r(\|x\|) \leqq r^0 \qquad (\|x\|^2 = x_1^2 + \cdots + x_n^2) \tag{5.3}$$

under (5.2). *Here* $r(u)$ *is independent of* $i, j_1, \cdots, j_n$ *and* r^0 *is independent of* $x_1, \cdots, x_n$; *furthermore*

$$\lim r(u) = 0 \qquad (as\ u \to 0); \tag{5.3a}$$

r^0 *may depend on* t, *in which case it is to be continuous for* $t \geqq t_0$.

Such systems (5.1) will include, in particular, the case when the $f_i(t; x_1, \cdots, x_n)$ are analytic in $x_1, \cdots, x_n$, for $|x_j| \leqq H$ and $t \geqq t_0$, and are continuous in t, the variables involved being subject to (5.2); $f_i(t; 0, \cdots, 0) = 0$. In the case when the term "analytic in $x_1, \cdots, x_n$," above, is replaced by the expression "analytic in $x_1, \cdots, x_n$ uniformly with respect to t $(t \geqq t_0)$"(7), one may choose $r(\|x\|)$ and r^0 in (5.3) independent of t.

It is observed that the system

$$\frac{dx_i}{dt} = l_i \qquad (i = 1, \cdots, n)$$

(7) In the sense of Liapounoff.

is precisely the previously investigated linear differential system (3.1).

Let $p_0(t)$ be the "least" function, continuous for $t \geqq t_0$ and such that

$$|p_{i,j}(t)| \leqq p_0(t) \tag{5.4}$$

and

$$|p_i^{j_1,\cdots,j_n}(t)| + r^0 \leqq p_0(t). \tag{5.4a}$$

Inasmuch as the number of continuous functions in the first members of (5.4), (5.4a) is finite, $p_0(t)$ may always be chosen as stated above.

The transformation

$$\tau = \phi(t) = \int_{t_0}^{t} p_0(t)dt \qquad (0 \leqq \tau \leqq \phi(+\infty) = \tau') \tag{5.5}$$

has the properties indicated at the beginning of §3. Under (5.5) the system (5.1) takes the form

$$\frac{dx_i}{d\tau} = \bar{f}_i = \bar{l}_i + \bar{q}_i \qquad (i = 1, \cdots, n),$$

where

$$\bar{l}_i = h_{1,i}(\tau)x_1 + \cdots + h_{n,i}(\tau)x_n,$$
$$\bar{q}_i = \sum_{j_1,\cdots} h_i^{j_1,\cdots,j_n}(\tau; x)x_1^{j_1} \cdots x_n^{j_n} \qquad (j_1 + \cdots + j_n = 2).$$

Here

$$h_{i,j}(\tau) = \frac{p_{i,j}(t)}{p_0(t)}, \qquad h_i^{j_1,\cdots,j_n}(\tau; x) = h_i^{j_1,\cdots,j_n}(\tau) + \bar{r}_i^{j_1,\cdots,j_n}(\tau; x),$$

with

$$h_i^{j_1,\cdots,j_n}(\tau) = \frac{1}{p_0(t)} p_i^{j_1,\cdots,j_n}(t), \qquad \bar{r}_i^{j_1,\cdots,j_n}(\tau; x) = \frac{1}{p_0(t)} r_i^{j_1,\cdots,j_n}(t; x).$$

In consequence of (5.3) and (5.4a)

$$|\bar{r}_i^{j_1,\cdots,j_n}(\tau; x)| \leqq \frac{1}{p_0(t)} r(\|x\|) = \bar{r}(\|x\|) \leqq \frac{r^0}{p_0(t)} = \bar{r}^0 \leqq 1, \tag{5.6}$$

$$|h_i^{j_1,\cdots,j_n}(\tau; x)| \leqq 1, \qquad |h_{i,j}(\tau)| \leqq 1, \qquad |h_i^{j_1,\cdots,j_n}(\tau)| \leqq 1 \tag{5.6a}$$

for $0 \leqq \tau \leqq \tau'$ and $|x_j| \leqq H$ $(j = 1, \cdots, n)$. It is noted that the function $\bar{r}^0$, in (5.6), is continuous in τ for $0 \leqq \tau \leqq \tau'$.

When the integral

$$\int_{t_0}^{\infty} p_0(t)dt \tag{5.7}$$

diverges, the interval for τ is $(0, +\infty)$; in the contrary case it is $(0, \tau')$, where $\tau' < +\infty$. In the latter case

$$p_0(t) \leqq b < \infty$$

and, in view of our purposes, the transformation (5.5) is not needed.

In view of the transformation (5.5) ($p_0(t)$ *from* (5.4), (5.4a)) *there is no loss of generality, if the general nonlinear problem* (5.1)–(5.1c) *is taken subject to the inequalities*

$$|p_{i,j}(t)|, \quad |p_i^{j_1,\cdots,j_n}(t)| \leqq b < \infty \qquad (j_1 + \cdots + j_n = 2), \tag{5.8}$$

$$|p_i^{j_1,\cdots,j_n}(t; x)| \leqq b, \tag{5.8a}$$

$$|r_i^{j_1,\cdots,j_n}(t; x)| \leqq r(\|x\|) \leqq r^0 \leqq b \tag{5.8b}$$

satisfied for $|x_j| \leqq H$ $(j=1, \cdots, n)$ *and* $t_0 \leqq t < \infty$, b *being independent of* t.

To the system (5.1), under (5.8)–(5.8b), we apply the transformation

$$z_i = x_i e^{\lambda t} \qquad (\lambda \text{ indeterminate}). \tag{5.9}$$

The system satisfied by the z_i will be of the form

$$\frac{dz_i}{dt} = p_{1,i}(t)z_1 + \cdots + (p_{i,i}(t) + \lambda)z_i + \cdots + p_{n,i}(t)z_n + p'_{n,i}$$

$(i=1, \cdots, n)$, with

$$p'_{n,i} = e^{-\lambda t} \sum_{\nu,\kappa}^{n} p_i^{\nu,\kappa} z_\nu z_\kappa, \qquad \delta p_i^{\nu,\kappa} = p_i^{j_1,\cdots,j_n}(t) + r_i^{j_1,\cdots,j_n}(t; x_1, \cdots, x_n) \tag{5.10}$$

$(j_1 + \cdots + j_n = 2)$, where sets (ν, κ) are in a certain correspondence with sets $(j_1, \cdots, j_n)$, while $\delta = 1$ for $\nu = \kappa$ and $\delta = 2$ for $\nu \neq \kappa$. We write

$$u = z_1^2 + \cdots + z_n^2. \tag{5.11}$$

Then

$$\frac{1}{2}\frac{du}{dt} = A(\lambda) + B(\lambda) + C(\lambda), \tag{5.12}$$

where $A(\lambda)$, $B(\lambda)$ are given by (3.9a) and

$$C(\lambda) = \sum_{i=1}^{n} z_i p'_{n,i}. \tag{5.12a}$$

Using the inequalities (5.8), from the developments of §3 we infer that for all $t \geqq t_0$.

$$(5.13)\qquad \begin{aligned} A(\lambda)+B(\lambda) &\geqq \eta u && (\text{all } \lambda \geqq (2n-1)b+\eta;\ \text{any } \eta>0),\\ A(\lambda)+B(\lambda) &\leqq -\eta u && (\text{all } \lambda \leqq -(2n-1)b-\eta)\end{aligned}$$

for all $t \geqq t_0$.

We think of the set of functions

$$x_1, \cdots, x_n$$

as a solution of (5.1); accordingly

$$(5.14)\qquad |x_j| \leqq H \qquad (j=1,\cdots,n;\ t \geqq t_0).$$

In consequence of these inequalities of (5.10) and of (5.8a)

$$|p_i^{\nu,\kappa}| \leqq b.$$

Hence by the inequality of Schwarz and by (5.10)

$$|p'_{n,i}| \leqq e^{-\lambda t} nbu \qquad (\text{cf. (5.11)}).$$

Accordingly, the function $C(\lambda)$ of (5.12a) will satisfy

$$|C(\lambda)|^2 \leqq \sum_i z_i^2 \sum_i (p'_{n,i})^2 \leqq b^2 e^{-2\lambda t} n^3 u^3$$

for all real λ and for all $t \geqq t_0$. Now, by (5.14), (5.9) and (5.11)

$$u^{1/2} \leqq n^{1/2} H e^{\lambda t};$$

whence

$$(5.15)\qquad |C(\lambda)| \leqq Hn^2 bu.$$

On choosing η of (5.13) as

$$\eta = Hn^2 b + \epsilon \qquad (\epsilon > 0),$$

by virtue of (5.12), (5.13) and (5.15) it is concluded that

$$\frac{1}{2}\frac{du}{dt} \geqq \epsilon u \qquad (\text{all } \lambda \geqq b(Hn^2+2n-1)+\epsilon = \lambda_0).$$

Similarly, it is shown that

$$\frac{1}{2}\frac{du}{dt} \leqq -\epsilon u \qquad (\text{all } \lambda \leqq -\lambda_0).$$

Therefore

$$u \geqq ce^{2\epsilon t} \quad (\text{all } \lambda \geqq \lambda_0), \qquad u \leqq c_1 e^{-2\epsilon t} \quad (\text{all } \lambda \leqq -\lambda_0)$$

for all $t \geqq t_0$; here c, c_1 are positive constants. Now ϵ (>0) may be taken arbi-

trarily small. Thus, *on writing* $v=x_1^2+\cdots+x_n^2$, *we find that*

$$\lim_t ve^{2\lambda t} = +\infty \qquad (\textit{all } \lambda > b(Hn^2+2n-1)),$$
$$\lim_t ve^{2\lambda t} = 0 \qquad (\textit{all } \lambda < -b(Hn^2+2n-1)). \tag{5.16}$$

With the aid of reasoning of the type employed subsequent to (3.8), from (5.16) it is inferred that

$$|C(x_1^2+\cdots+x_n^2|e^t)| \leqq 2b(Hn^2+2n-1), \tag{5.17}$$

and

$$|C(x_1,\cdots,x_n|e^t)| \leqq b(Hn^2+2n-1). \tag{5.17a}$$

In view of (5.14) it is clear that

$$C(x_1^2+\cdots+x_n^2|e^t) \geqq 0.$$

THEOREM 5.1. *Consider the nonlinear differential system, as described in connection with* (5.1)–(5.3a). *In view of the statement with reference to* (5.8)–(5.8b) (*note transformation* (5.5)) *there is no loss of generality in assuming* (5.8)–(5.8b). *If a set of functions*

$$x_1,\cdots,x_n$$

satisfies the system for all $t\geqq t_0$, *so that*

$$|x_j| \leqq H \qquad (j=1,\cdots,n),$$

then necessarily

$$0 \leqq C(x_1^2+\cdots+x_n^2|e^t) \leqq 2b(Hn^2+2n-1).$$

More precisely, (5.16) *will hold.*

It is observed that the information supplied by the inequality, last displayed in the above theorem, amounts to an assertion to the effect that $x_1^2+\cdots+x_n^2$ cannot approach zero (as $t\to\infty$) faster than at a certain rate.

Consider the case when, for $t\to\infty$,

$$\lim p_{i,j}(t) = \lim p_i^{j_1,\cdots,j_n}(t) = \lim r^0 = 0. \tag{5.18}$$

The function $p_0(t)$ of (5.4), (5.4a) will then tend to zero with $1/t$. There exists a "least" monotone non-increasing function $m(t)$, continuous for $t\geqq t_0$, for which

$$p(t) \leqq m(t), \qquad \lim m(t) = 0. \tag{5.18a}$$

If one had $m(t_1)=0$ for some $t_1 \geqq t_0$ we would have

$$f_i(t; x_1, \cdots, x_n) \equiv 0 \quad (\text{all } t \geqq t_1;\ i = 1, \cdots, n);$$

the system (5.1) would then be trivial from our point of view. Hence we may take

$$m(t) > 0. \tag{5.18b}$$

In (5.16), on letting $v = x_1^2 + \cdots + x_n^2$, we have

$$\begin{aligned} &\lim_t v \exp\left\{[2(Hn^2 + 2n - 1)m(\tau) + \epsilon/2]t\right\} = +\infty, \\ &\lim_t v \exp\left\{[-2(Hn^2 + 2n - 1)m(\tau) - \epsilon/2]t\right\} = 0 \end{aligned} \tag{5.19}$$

for all $\epsilon > 0$ and for all $\tau \geqq \tau_0$. For any ϵ, such that

$$0 < \epsilon \leqq \epsilon_0 < 4m(t_0)(Hn^2 + 2n - 1),$$

the equation

$$2(Hn^2 + 2n - 1)m(\tau) = \epsilon/2$$

will have a solution $\tau = \tau_\epsilon \geqq t_0$. Substitution of $\tau = \tau_\epsilon$ in (5.19) will yield

$$\lim_t v e^{\epsilon t} = \infty, \qquad \lim_t v e^{-\epsilon t} = 0;$$

that is,

$$C_\epsilon(x_1^2 + \cdots + x_n^2 \mid e^t) = 0. \tag{5.20}$$

We have established

THEOREM 5.2. *Consider the nonlinear differential system* (5.1). *Suppose that the functions, involved, vanish for* $t \to \infty$, *as indicated in* (5.18). *Then for every solution*

$$x_1, \cdots, x_n$$

of (5.1), *defined for* $t \geqq t_0$, *we shall have* (5.20).

With the aid of a reasoning of the type used subsequent to Theorem 3.2 we obtain the following corollary to Theorem 5.2.

Consider the system (5.1). Let $p_0(t)$ be the function of (5.4), (5.4a). If the integral

$$\int_{t_0}^{\infty} p_0(t)dt \tag{5.21}$$

diverges, then for every solution $x_1, \cdots, x_n$ of (5.1) (with $|x_i| \leqq H$) we shall have

$$C_*\left(x_1^2 + \cdots + x_n^2 \,\middle|\, \exp \int_{t_0}^{t} \rho(t)dt\right) = 0 \tag{5.22}$$

for every continuous function $\rho(t)$ $(\geqq 0)$ such that

$$\lim_{t\to\infty} \frac{p_0(t)}{\rho(t)} = 0.$$

If the integral (5.21) converges (or, more generally, if $\lim p_0(t) = 0$), then for every solution $x_1, \cdots, x_n$ one has (5.22) for every continuous function $\rho(t)$ $(\geqq 0)$ for which

$$\lim \frac{p_0(t)}{\rho(t)} = 0, \qquad \int_{t_0}^{\infty} \rho(t)dt \text{ diverges.}$$

6. **Method of successive approximations.** In this section we shall establish certain "growth" properties for solutions of the nonlinear system (5.1), making use of the solutions of the corresponding linear problem

$$\frac{dx_j}{dt} = l_j(t; x) = p_{1,j}(t)x_1 + \cdots + p_{n,j}(t)x_n \qquad (j = 1, \cdots, n). \tag{6.1}$$

We shall use solutions of (6.1) as the first approximation.

For some continuous function $p^0(t)$ we have

$$\left| p_j^{j_1,\cdots,j_n}(t) \right| \leqq p^0(t). \tag{6.2}$$

We shall write

$$q_j = q_j(t; x) = \alpha_j(t; x_1, \cdots, x_n) + \beta_j(t; x_1, \cdots, x_n), \tag{6.3}$$

$$\alpha_j(t; x_1, \cdots, x_n) = \sum_{j_1+\cdots+j_n=2} p_j^{j_1,\cdots,j_n}(t)x_1^{j_1} \cdots x_n^{j_n} = \sum_{\nu,m=1}^{n} p_j^{\nu,m}(t)x_\nu x_m, \tag{6.3a}$$

$$\beta_j(t; x_1, \cdots, x_n) = \sum_{j_1+\cdots+j_n=2} r_j^{j_1,\cdots,j_n}(t; x)x_1^{j_1} \cdots x_n^{j_n} = \sum_{\nu,m=1}^{n} r_j^{\nu,m}(t; x)x_\nu x_m; \tag{6.3b}$$

here

$$p_j^{\nu,m}(t) = p_j^{m,\nu}(t), \qquad r_j^{\nu,m}(t; x) = r_j^{m,\nu}(t; x). \tag{6.3c}$$

In order to carry out the method of successive approximations a special continuity condition for the functions

$$r_j^{\nu,m}(t; x) = r_j^{\nu,m}(t; x_1, \cdots, x_m) \tag{6.4}$$

is needed. It will be convenient to introduce a hypothesis in the nature of a *Lipschitz* condition.

HYPOTHESIS 6.5. *The functions* (6.4), *involved in* (6.3b), *satisfy inequalities*

$$(6.5a)\quad \left| r_j^{r,m}(t; z_1 + x_1, \cdots, z_n + x_n) - r_j^{r,m}(t; z_1, \cdots, z_n) \right| \leqq r_1(t)x^* \quad (t \geqq t_0),$$

whenever

$$(6.5b)\quad |z_j| \leqq z^*, \qquad |x_j| \leqq x^* \quad (j = 1, \cdots, n), \qquad z^* + x^* \leqq H;$$

here $r_1(t)$ *is continuous.*

The following lemma will be stated without proof.

LEMMA 6.1. *Under Hypothesis* 6.5 *the functions* $q_j = q_j(t; x)$, *involved in the system* (5.1), *satisfy the inequalities*

$$(6.6)\quad \left| q_j(t; z_1 + x_1, \cdots, z_n + x_n) - q_j(t; z_1, \cdots, z_n) \right| \leqq q_0(t)(z^* + x^*)x^*$$

$(j = 1, \cdots, n; t \geqq t_0)$, *where*

$$(6.6a)\quad q_0(t) = 2n^2(p^0(t) + r^0(t)) + n^2Hr_1(t)$$

[*cf.* (6.2), (5.3) $(r^0 = r^0(t))$, (6.5a)], *whenever* (6.5b) *holds.*

We shall now prove the following result.

LEMMA 6.2. *Effecting on* t, *if necessary, a suitable transformation*

$$(6.7)\quad \tau = \int_{t_0}^{t} q'(t)dt \qquad (\text{continuous } q'(t) > 0),$$

the function $q_0(t)$, *involved in* (6.6), *may be replaced by a constant* q_0. *Moreover, it can be arranged at the same time to have the* $|p_{i,j}(t)|$ *of* (6.1) *uniformly bounded.*

In fact, if $q_0(t)$ is uniformly bounded we may define q_0 as the least upper bound of $q_0(t)$. In the contrary case the integral

$$\int_{t_0}^{\infty} q_0(t)dt$$

will diverge; the transformation (6.7), with $q'(t) = q_0(t)$, will yield the system

$$\frac{dx_j}{d\tau} = \frac{1}{q_0(t)} l_j(t; x) + \frac{1}{q_0(t)} q_j(t; x) = \bar{l}_j(\tau; x) + \bar{q}_j(\tau; x),$$

for which the function $\bar{q}_0(\tau)$, corresponding to $q_0(t)$, is given by

$$\bar{q}_0(\tau) = 1 \qquad (0 \leqq \tau < \infty).$$

If it were desired to have $q_0(t)$ in (6.6) replaced by a constant and at the same time to have the $|p_{i,j}(t)|$, in (6.1), uniformly bounded, we would proceed as follows. Let $p(t)$ be the "least" continuous function such that

$$|p_{i,j}(t)| \leqq p(t).$$

If $q_0(t)$ is not uniformly bounded, let $q'(t)$ be the "least" continuous function such that

$$q_0(t), \qquad p(t) \leqq q'(t).$$

Then the integral

$$\int_{t_0}^{\infty} q'(t)dt \tag{6.8}$$

will be divergent and the transformation (6.7) will produce the desired effect. If $q_0(t)$ is uniformly bounded, but $p(t)$ is not, then we still define $q'(t)$ as above. The integral (6.8) will be divergent again. The transformation (6.7) will serve the purpose.

The developments in the sequel will be under the following hypothesis relating to the linear problem (6.1).

HYPOTHESIS 6.9. *There exists a matrix solution* $X(t)=(x_{i,j}(t))$ $(i, j=1, \cdots, n)$ *of the matrix equation*

$$X^{(1)}(t) = X(t)P(t), \qquad P(t) = (p_{i,j}(t)) \tag{6.9a}$$

such that

$$|x_{i,j}(t)| \leqq \phi_i(t) \qquad (i = 1, \cdots, n; t \geqq t_0), \tag{6.9b}$$

while the elements $\bar{x}_{i,j}(t)$ *from the matrix* $X^{-1}(t)=(\bar{x}_{i,j}(t))$ *satisfy*

$$|\bar{x}_{i,j}(t)| \leqq \frac{1}{\phi_j(t)}\bar{\phi}(t). \tag{6.9c}$$

Here the $\phi_i(t)$, $\bar{\phi}(t)$ *are continuous. There exists a continuous function* $\theta(t)$ *so that*

$$\phi_1(t) \leqq \phi_2(t) \leqq \cdots \leqq \phi_m(t) \leqq \cdots \leqq \phi_n(t), \tag{6.9d}$$

$$\lim \phi_j(t) = 0 \qquad (j = 1, \cdots, m), \tag{6.9d$_1$}$$

$$\bar{\phi}(t)\theta(t) \leqq b, \qquad \frac{1}{\theta(t)}\phi_m(t)(t-t') \leqq b_0 \qquad (t \geqq t_0; \text{ some } t' < t_0), \tag{6.9d$_2$}$$

$$\frac{\phi_m^2(t)(t-t')1+\epsilon'}{\theta(t)\phi_1(t)} \to 0 \quad [\text{monotonically as } t\,(\geqq t_0) \to +\infty; \text{ some } \epsilon' > 0], \tag{6.9d$_3$}$$

$$\lim \frac{\phi_m(t)}{\theta(t)}(t-t') = 0. \tag{6.9d$_4$}$$

Moreover, the functions $\phi_1(t)\phi_\lambda^{-1}(t)$ $(\lambda=1, \cdots, n)$ *are monotone non-increasing.*

Further along in the text an important example of linear differential systems will be given for which conditions of Hypothesis 6.9 hold.

If $C=(c_{i,j})$ is a matrix of arbitrary constants with determinant $|C|$, distinct from zero, then

$$CX(t) = \left(\sum_{\tau} c_{i,\tau} x_{\tau,j}(t) \right)$$

will constitute a matrix solution of (6.9a). With (6.9b), (6.9d$_1$) in view, we form functions

$$x_{j;0}(t) = \sum_{\tau=1}^{m} c_{i,\tau} x_{\tau,j}(t) \qquad (j = 1, \cdots, n);$$

for a fixed i these functions will constitute a solution of (6.9a). Accordingly

$$x_{j;0}(t) = c_1 x_{1,j}(t) + \cdots + c_m x_{m,j}(t) \qquad (j = 1, \cdots, n) \tag{6.10}$$

will constitute a solution of the linear system.

A solution of (5.1) will be obtained in the form of series

$$x_j(t) = x_{j;0}(t) + x_{j;1}(t) + \cdots + x_{j;\nu}(t) + \cdots \qquad (j = 1, \cdots, n), \tag{6.11}$$

where

$$\frac{dx_{j;0}}{dt} = l_j(t; x_{;0}) \qquad (j = 1, \cdots, n),$$

$$\frac{dx_{j;\nu}}{dt} = l_j(t; x_{;\nu}) + g_{j,\nu}(t) \qquad (j = 1, \cdots, n; \nu = 1, 2, \cdots);$$

here

$$g_{j,\nu}(t) = q_j(t; x_{;\nu-1} + z_{;\nu-2}) - q_j(t; z_{;\nu-2}) \qquad (\nu = 1, 2, \cdots), \tag{6.12}$$

where $z_{;-1}=0$ and

$$z_{j;\nu} = x_{j;0} + x_{j;1} + \cdots + x_{j;\nu} \quad (\nu \geqq 0), \qquad q_j(t; z_{;-1}) = 0.$$

The notation above, for instance for $l_i(t; x_{;\nu})$, is to be understood in the sense that

$$l_j(t; x_{;\nu}) = l_j(t; x_{1;\nu}, \cdots, x_{n;\nu}).$$

If, with $\nu \geqq 1$, $g_{j,\nu}(t)$ is thought of as known, the $x_{j;\nu}$ $(j=1, \cdots, n)$ will be given by

$$x_{j;\nu}(t) = \sum_{\lambda=1}^{n} w^{\lambda,\nu}(t) x_{\lambda,j}(t) \qquad (j = 1, \cdots, n), \tag{6.13}$$

where

$$w^{\lambda,\nu}(t) = \sum_{\tau=1}^{n} \int^{t} g_{\tau,\nu}(u) \bar{x}_{\tau,\lambda}(u) du \qquad (\lambda = 1, \cdots, n), \tag{6.13a}$$

provided that the integrals here can be evaluated(8). In determinant form the $x_{\tau,\lambda}$ are given by

$$(6.13b)\qquad (-1)^{\tau+\lambda}\Delta x_{\tau,\lambda} = \begin{vmatrix} x_{1,1}, & \cdots, x_{1,\tau-1}, & x_{1,\tau+1}, & \cdots, x_{1,n} \\ \cdot & \cdots \cdot & \cdot & \cdots \cdot \\ x_{\lambda-1,1}, & \cdots \cdot & \cdot & \cdots \cdot \\ x_{\lambda+1,1}, & \cdots \cdot & \cdot & \cdots \cdot \\ \cdot & \cdots \cdot & \cdot & \cdots \cdot \\ x_{n,1}, & \cdots \cdot & \cdot & \cdots \cdot \end{vmatrix},$$

where Δ denotes the determinant

$$(6.13c)\qquad \Delta = |(x_{i,j}(t))| = \gamma \exp \int^t (p_{1,1}(u) + \cdots + p_{n,n}(u))du.$$

The constants $c_1, \cdots, c_m$ of (6.10) will be left arbitrary, except for the inequalities

$$(6.14)\qquad |c_j| \leqq c_0 \qquad (j = 1, \cdots, m;\ c_0 > 0).$$

We take c_0 sufficiently small and t_0 sufficiently great so that

$$(6.15)\qquad \begin{gathered} x_0^*(t) = mc_0\phi_m(t) \leqq \tfrac{1}{2}H, \\ q_0[mc_0 + 2nn_0c_0^2\theta^{-1}(t)\phi_m(t)(t-t')] \leqq 1 \qquad (n_0 = (n/\epsilon')bq_0m^2), \end{gathered}$$

and so that

$$(6.15a)\qquad x_1^*(t) = nn_0c_0^2\theta^{-1}(t)\phi_m^2(t)(t-t') \leqq 2^{-2}H,$$

$$(6.15b)\qquad \frac{n^2}{\epsilon'}bb_0 \leqq \tfrac{1}{2} \qquad \text{(cf. (6.9d}_2\text{))}.$$

In this connection q_0 is a constant, which according to Lemma 6.2 can be used in (6.6) in place of $q_0(t)$; (6.15b) may be satisfied on noting that, in view of (6.9d$_4$), the constant b_0 of (6.9d$_2$) can be made arbitrarily small by taking t_0 sufficiently great.

By (6.10), (6.14), (6.9b) and (6.9d)

$$(6.16)\qquad |x_{j:0}(t)| \leqq mc_0\phi_m(t) = x_0^*(t) \qquad \text{(cf. (6.15))}.$$

In consequence of (6.12) and of Lemma 6.1 (with $q_0(t)$ replaced by a constant q_0)

$$|g_{i,1}(t)| = |q_i(t; x_{1:0}, \cdots, x_{n:0}) - q_i(t; 0, \cdots, 0)| \leqq q_0(x_0^*(t))^2,$$

(8) With the capital letters denoting matrices (of n^2 elements), it is observed that a matrix solution of $X^{(1)} = XP + G$ may be expressed in the form $X = WX_0$, where $X_0^{(1)} = X_0P$ and $W^{(1)} = GX_0^{-1}$. This will yield (6.13), (6.13a).

where $x_0^*(t)$ is the function of (6.16). Accordingly, by (6.9c)

$$|\, g_{\tau,1}(u)\bar{x}_{\tau,\lambda}(u)\,| \leqq q_0 m^2 c_0^2 \phi_m^2(u)\,\frac{1}{\phi_\lambda(u)}\,\bar{\phi}(u) \qquad (u \geqq t_0).$$

Thus, in consequence of (6.9d$_2$),

$$|\, g_{\tau,1}(u)\bar{x}_{\tau,\lambda}(u)\,| \leqq b q_0 m^2 c_0^2\,\frac{\phi_m^2(u)}{\theta(u)\phi_\lambda(u)}\,. \tag{6.17}$$

Now

$$I_\lambda(u) = \frac{\phi_m^2(u)(u-t')^{1+\epsilon'}}{\theta(u)\phi_\lambda(u)} = \frac{\phi_m^2(u)(u-t')^{1+\epsilon'}}{\theta(u)\phi_1(u)}\cdot\frac{\phi_1(u)}{\phi_\lambda(u)}\,. \tag{6.18}$$

The two factors in the second member above are monotone non-increasing for $u \geqq t'$, as follows from (6.9d$_3$) and by the statement subsequent to (6.9d$_4$); the first member in (6.18) will have the same property. Hence

$$I_\lambda(u) \leqq I_\lambda(t) \qquad \text{(for all } u \geqq t \geqq t_0\text{; cf. (6.18))}. \tag{6.19}$$

From (6.17) on using (6.19) we derive

$$\begin{aligned}\left|\int_\infty^t g_{\tau,1}(u)\bar{x}_{\tau,\lambda}(u)du\right| &\leqq \int_t^\infty b q_0 m^2 c_0^2\,\frac{\phi_m^2(u)}{\theta(u)\phi_\lambda(u)}\,du \\ &= b q_0 m^2 c_0^2 \int_t^\infty I_\lambda(u)(u-t')^{-1-\epsilon'}du \\ &\leqq b q_0 m^2 c_0^2 I_\lambda(t)\int_t^\infty (u-t')^{-1-\epsilon'}du \\ &= \frac{1}{\epsilon'}\,b q_0 m^2 c_0^2\,\frac{\phi_m^2(t)(t-t')}{\theta(t)\phi_\lambda(t)}\,.\end{aligned} \tag{6.20}$$

In consequence of (6.13a) and (6.20)

$$|\, w^{\lambda,1}(t)\,| \leqq \sum_{\tau=1}^n \left|\int_\infty^t g_{\tau,1}(u)\bar{x}_{\tau,\lambda}(u)du\right| \leqq n_0 c_0^2\,\frac{\phi_m^2(t)(t-t')}{\theta(t)\phi_\lambda(t)}\,;$$

here n_0 has the value assigned in (6.15). Hence, by virtue of (6.13) and (6.9b), one has

$$|\, x_{j;1}(t)\,| \leqq \sum_{\lambda=1}^n |\, w^{\lambda,1}(t)\,|\,|\, x_{\lambda,j}(t)\,| \leqq n n_0 c_0^2 \theta^{-1}(t)\phi_m^2(t)(t-t') = x_1^*(t);$$

in this connection, we note, (6.15a) is satisfied.

In view of (6.12)

$$g_{j,2}(t) = q_j(t;\, x_{:1} + z_{:0}) - q_j(t;\, z_{:0}) \qquad (z_{:0} = x_{:0}).$$

Thus, by Lemma 6.1,

$$\begin{aligned} |\, g_{j,2}(t)\,| &\leqq q_0[x_0^*(t) + x_1^*(t)]x_1^*(t) \\ &= q_0\left[mc_0 + nn_0c_0^2\frac{\phi_m(t)(t-t')}{\theta(t)}\right]\phi_m(t)x_1^*(t). \end{aligned}$$

Hence, in view of (6.15), $|\, g_{j,2}(t)\,| \leqq \phi_m(t)x_1^*(t)$.

Whence by (6.9c) and (6.9d$_2$)

$$\begin{aligned} |\, g_{\tau,2}(u)\bar{x}_{\tau,\lambda}(u)\,| &\leqq x_1^*(u)\frac{\phi_m(u)}{\phi_\lambda(u)}\bar{\phi}(u) \leqq bx_1^*(u)\frac{\phi_m(u)}{\theta(u)\phi_\lambda(u)} \\ &= bnn_0c_0^2\left\{\frac{\phi_m(u)(u-t')}{\theta(u)}\right\}\left\{\frac{\phi_m^2(u)(u-t')^{1+\epsilon'}}{\theta(u)\phi_\lambda(u)}\right\}(u-t')^{-1-\epsilon'} \end{aligned} \tag{6.21}$$

and, in consequence of (6.9d$_2$),

$$|\, g_{\tau,2}(u)\bar{x}_{\tau,\lambda}(u)\,| \leqq bb_0nn_0c_0^2I_\lambda(u)(u-t')^{-1-\epsilon'} \qquad \text{(cf. (6.18))}$$

for $u \geqq t_0$. This, together with (6.19), implies

$$\left|\int_\infty^t g_{\tau,2}(u)\bar{x}_{\tau,\lambda}(u)du\right| \leqq \frac{1}{\epsilon'}bb_0nn_0c_0^2I_\lambda(t)(t-t')^{-\epsilon'};$$

that is,

$$\left|\int_\infty^t g_{\tau,2}(u)\bar{x}_{\tau,\lambda}(u)du\right| \leqq \frac{1}{\epsilon'}bb_0nn_0c_0^2\frac{\phi_m^2(t)(t-t')}{\theta(t)\phi_\lambda(t)}.$$

Accordingly by (6.13a)

$$|\, w^{\lambda,2}(t)\,| \leqq nn_2c_0^2\frac{\phi_m^2(t)(t-t')}{\theta(t)\phi_\lambda(t)} \qquad (n_2 = (1/\epsilon')bb_0nn_0), \tag{6.22}$$

and, by virtue of (6.13),

$$|\, x_{j;2}(t)\,| \leqq \sum_{\lambda=1}^n |\, w^{\lambda,2}(t)\,|\,\phi_\lambda(t) \leqq n^2n_2c_0^2\frac{\phi_m^2(t)(t-t')}{\theta(t)} = x^*(2,t). \tag{6.23}$$

In view of the definition of n_2, given in (6.22), comparison of (6.23) and (6.15a) will yield

$$x^*(2,t) = \frac{n^2}{\epsilon'}bb_0x_1^*(t) \leqq x_2^*(t) = 2^{-1}x_1^*(t),$$

inasmuch as (6.15b) holds.

Suppose that for some $N>2$ we have

$$|x_{j;\nu}(t)| \leqq x_\nu^*(t) = 2^{-(\nu-1)}x_1^*(t) \qquad (\nu = 1, \cdots, N-1), \tag{6.24}$$

where $x_1^*(t)$ is from (6.15a). It will be established by induction that (6.24) takes place for $\nu = 1, 2, \cdots$.

In view of (6.24)

$$|z_{j;N-2}(t)| \leqq |x_{j;0}(t)| + \cdots + |x_{j;N-2}(t)| \leqq x_0^*(t) + x_1^*(t) + 2^{-1}x_1^*(t) + \cdots + 2^{-(N-3)}x_1^*(t).$$

For $N=2$ the last member above will consist of $x_0^*(t)$ only. Thus

$$|z_{j;N-2}(t)| \leqq x_0^*(t) + 2x_1^*(t) = z^*(t) \tag{6.25}$$

and, by virtue of (6.15), (6.15a), one has $z^*(t) \leqq H$.

From (6.12) (with $\nu = N$) and Lemma 6.1 it is concluded that

$$|g_{j;N}(t)| = |q_j(t; x_{;N-1} + z_{;N-2}) - q_j(t; z_{;N-2})| \leqq q_0 \cdot (z^* + x^*)x^*$$

here x^*, z^* are from the inequalities

$$|x_{j;N-1}(t)| \leqq x^* = x_{N-1}^*(t) = 2^{-(N-2)}x_1^*(t) \qquad \text{(cf. (6.24))},$$

$$|z_{j;N-2}(t)| \leqq z^* = z^*(t) \qquad \text{(cf. (6.25))}.$$

Thus, by (6.15) and (6.15b)

$$\begin{aligned} |g_{j,N}(t)| &\leqq q_0(x_0^*(t) + 2x_1^*(t))2^{-(N-2)}x_1^*(t) \\ &= q_0\left\{mc_0 + 2nn_0c_0^2\frac{\phi_m(t)(t-t')}{\theta(t)}\right\}\phi_m(t)2^{-(N-2)}x_1^*(t) \qquad (6.26) \\ &\leqq 2^{-(N-2)}\phi_m(t)x_1^*(t). \end{aligned}$$

In view of (6.26) and (6.9c)

$$|g_{\tau,N}(u)\bar{x}_{\tau,\lambda}(u)| \leqq 2^{-(N-2)}x_1^*(u)\frac{\phi_m(u)}{\phi_\lambda(u)}\bar{\phi}(u).$$

The second member, here, is equal to the second member of (6.21), multiplied by $2^{-(N-2)}$. Carrying out the steps from (6.21) to (6.22), with the indicated modification, from (6.13a) it is deduced that

$$\left|\int_\infty^t g_{\tau,N}(u)\bar{x}_{\tau,\lambda}(u)du\right| \leqq 2^{-(N-2)}nn_2c_0^2\frac{\phi_m^2(t)(t-t')}{\theta(t)\phi_\lambda(t)}. \tag{6.27}$$

With the aid of (6.13) (for $\nu = N$) from (6.27) one finds

$$|x_{j;N}(t)| \leqq 2^{-(N-2)}n^2n_2^2c_0^2\frac{\phi_m^2(t)(t-t')}{\theta(t)}.$$

Substituting the expression n_3 from (6.22) and making use of (6.15b), one obtains

$$| x_{j;N}(t) | \leqq 2^{-(N-1)} n n_0 c_0^2 \frac{\phi_m^2(t)(t - t')}{\theta(t)} = 2^{-(N-1)} x_1^*(t) = x_N^*(t)$$

for $t \geqq t_0$.

Thus, *by induction it has been established that* (6.24) *holds for all* $\nu = 1, 2, \cdots$.

Turning now to the series (6.11) and recapitulating the inequalities

$$| x_{j;0}(t) | \leqq m c_0 \phi_m(t) \leqq 2^{-1} H, \quad | x_{j;\nu}(t) | \leqq 2^{-(\nu-1)} x_1^*(t) \qquad (\nu = 1, 2, \cdots), \tag{6.28}$$

where

$$x_1^*(t) = \bar{n} c_0^2 \frac{\phi_m^2(t)(t - t')}{\theta(t)} \leqq 2^{-2} H \qquad (\bar{n} = n n_0), \tag{6.28a}$$

we obtain

$$| x_j(t) - x_{j;0}(t) | \leqq \sum_{\nu=1}^{\infty} 2^{-(\nu-1)} x_1^*(t) = 2 x_1^*(t) \qquad (\leqq \tfrac{1}{2} H). \tag{6.29}$$

THEOREM 6.1. *Consider the nonlinear differential system* (5.1) [*cf.* (6.1), (5.1b), (5.1c), (5.3), (6.2)] *under Hypotheses* 6.5, 6.9; *it is supposed that on making use of Lemma* 6.2, *the function* $q_0(t)$ (*cf.* (6.6a)) *involved in* (6.6) *has been replaced by a constant* q_0. *The series* (6.11), *with integrations suitably performed, will be convergent for* $t \geqq t_0$ *and will represent functions* $x_j(t)$ $(j = 1, \cdots, n)$ *constituting a solution of* (5.1). *This solution is approximated by the solution*

$$x_{j;0}(t) \qquad (j = 1, \cdots, n; \text{ cf. } (6.10))$$

of the linear problem. The approximation is in the sense indicated in (6.29) [(6.28), (6.28a)]. *The*

$$x_j(t) \qquad (j = 1, \cdots, n)$$

contain m *arbitrary constants* $c_1, \cdots, c_m$, *subject to inequalities*

$$| c_i | \leqq c_0 \qquad (i = 1, \cdots, m; c_0 > 0).$$

The numbers c_0, t_0 *are to be chosen according to the italicized statement in connection with* (6.15)–(6.15b).

Note. The above developments hold in their essential features when complex values are admitted and t is restricted to a ray. It is also observed that results of the type given in Theorem 6.1 will continue to hold when the hypotheses on the $q_i(t; x)$ (in (5.1b)) are somewhat lightened; in establishing

Theorem 6.1 we have not used the fact that

$$\lim r_j^{j_1,\cdots,j_n}(t; x) = 0 \qquad (\text{as } x_1^2 + \cdots + x_n^2 \to 0).$$

The hypotheses originally given, however, are necessary for certain significant further investigations regarding some questions of interest in dynamics.

With a view to an important application of Theorem 6.1 let us consider the following case. After reducing $q_0(t)$ in (6.6a) to a constant q_0, using if necessary a transformation indicated in Lemma 6.2, suppose that the linear problem (6.1) presents the classical "irregular singular point" problem, with the singular point at $t=\infty$; that is, suppose that the $p_{i,j}(t)$ in (6.1) are functions, which at least on the interval $(t_0, +\infty)$ are asymptotic (at $t=+\infty$) to series (possibly divergent) of the form

$$x^{\mu'/\kappa}[a_0 + a_1 x^{-1/\kappa} + \cdots + a_\sigma x^{-\sigma/\kappa} + \cdots] \qquad (\text{integers } \kappa\ (>0),\ \mu').$$

This, of course, would include constants, polynomials, rational functions, as well as functions analytic in the complex t-plane for $|t| \geqq r_0 > 0$ and, possibly, having a pole at $t=\infty$. With regard to the analytic theory of the problem of the "irregular singular point" (for linear differential equations) it is to be said that a complete solution from various points of view has been given by *W. J. Trjitzinsky*(9).

On taking account of the developments given in (T) and of (6.13b), (6.13c) it may be shown that, except for (6.9d$_1$)–(6.9d$_4$), all the conditions of Hypothesis 6.9 can be made to hold with the following meaning of the symbols involved. We take t_0 (>0) sufficiently great; the $\phi_i(t)$ are certain functions of the form

$$\phi_i(t) = \{\exp[\phi_{i,\mu}t^{\mu/p} + \phi_{i,\mu-1}t^{(\mu-1)/p} + \cdots + \phi_{i,1}t^{1/p}]\}t^{\tau_i}\gamma'$$

(6.30) $[\gamma' > 0;\ \tau_i, \phi_{i,\nu}$ real; integers $\mu\ (>0)$, $p\ (>0)$; $p =$ integral multiple of $\kappa]$;

$\bar{\phi}(t)$ is given by

(6.30a) $$\bar{\phi}(t) = \gamma'' t^\alpha \qquad (\text{some } \alpha \geqq 0;\ \gamma'' > 0).$$

With t_0 suitably great and the $\phi_i(t)$ suitably ordered, not only will conditions (6.9b), (6.9c) be satisfied for a certain matrix solution $X(t)$ of (6.9a), but also (6.9d) as well as the condition stated subsequent to (6.9d$_4$) will be satisfied. Now the remaining relations (6.9d$_1$)–(6.9d$_4$) will not hold, with

(6.31) $$t' = 0, \qquad \theta(t) = t^{-\alpha},$$

(9) W. J. Trjitzinsky, *Analytic theory of linear differential equations*, Acta Mathematica, vol. 62 (1934), pp. 167–226, in the sequel referred to as (T).

W. J. Trjitzinsky, *Laplace integrals and factorial series in the theory of linear differential and linear difference equations*, these Transactions, vol. 37 (1935), pp. 80–146.

in every case; however, $(6.9d_1)$–$(6.9d_4)$ will take place in an extensive variety of cases. By (6.30)

$$\phi_i(t) = \gamma^i t^{r_i} \exp\left[t^{\mu/p}(\phi_{i,\mu} + t^{-1/p}\beta_i(t))\right],$$

where the $|\beta_i(t)|$ are uniformly bounded. Having arranged (6.9d), necessarily

$$\phi_{1,\mu} \leqq \phi_{2,\mu} \leqq \cdots \leqq \phi_{n,\mu},$$

if we have $\mu > 0$. To insure $(6.9d_1)$ it suffices to have

$$\phi_{1,\mu} \leqq \cdots \leqq \phi_{m,\mu} < 0. \tag{6.32}$$

Under (6.32) conditions $(6.9d_1)$, $(6.9d_4)$ will hold. Finally, $(6.9d_3)$ will be secured (with t_0 sufficiently great) provided that in addition to (6.32) one has

$$2\phi_{m,\mu} < \phi_{1,\mu}. \tag{6.33}$$

The truth of the latter statement follows from the fact that

$$t^{2r_m - r_1 + \alpha + 1 + s'} \exp\left\{t^{\mu/p}\left[(2\phi_{m,\mu} - \phi_{1,\mu}) + t^{1/p}\beta_{1,m}(t)\right]\right\}$$

is equal to the first member in $(6.9d_3)$, except for a constant factor, $|\beta_{1,m}(t)|$ being uniformly bounded.

It is observed that, *if* (6.32) *holds for some* $m \geqq 1$, *we shall have* (6.33) *satisfied at least for* $m = 1$. *Accordingly, it is noted that if in* (6.30) *at least one* $\phi_{i,\mu}$, *say* $\phi_{1,\mu}$, *is negative (while* $\mu/p > 0$), *all the conditions of Hypothesis* 6.9 *will be satisfied (with functions* (6.30)–(6.31)) *on the interval*

$$(t_0, +\infty) \qquad (t_0 \textit{ suitably great}).$$

With (6.32), (6.33) valid for some $m \geqq 1$, the solution

$$x_j(t) \qquad (j = 1, \cdots, n)$$

of the nonlinear problem (5.1), referred to in Theorem 6.1, will be approximated by the solution $x_{j;0}(t)$ $(j = 1, \cdots, n)$ of the linear problem as follows:

$$\begin{aligned} |x_j(t) - x_{j;0}(t)| &\leqq 2\bar{n}c_0^2\phi_m^2(t)t^{1+\alpha} \\ &= \bar{n}_0 c_0^2 t^{2r_m} \exp\left\{2t^{\mu/p}\left[\phi_{m,\mu} + t^{-1/p}\beta_m(t)\right]\right\} \\ &\qquad [t \geqq t_0;\ |\beta_m(t)| \leqq \beta_m < \infty;\ \text{constant } \bar{n}_0 > 0;\ \phi_{m,\mu} < 0]; \end{aligned}$$

here

$$|x_{j;0}(t)| \leqq mc_0\phi_m(t) \qquad (j = 1, \cdots, n;\ \text{cf. (6.30)}).$$

If $\mu/p > 0$, *approximation results of essentially the above type are obtained in all cases for intervals, extending to infinity, along suitable rays*

$$\text{angle of } t = \text{constant},$$

provided that on those intervals the nonlinear component of the problem is of the type which we specified above.

We shall have $\mu/p>0$, whenever the singularity, at $t=\infty$, of the linear problem is formally irregular.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.,
THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

NORMAL ALGEBRAIC NUMBER FIELDS

BY

SAUNDERS MAC LANE AND O. F. G. SCHILLING

Introduction. In this paper we present a detailed account of the results recently published in the Proceedings of the National Academy of Sciences [29](1). Our theory is an attempt to generalize the results of the classical class field theory to arbitrary normal fields. In the last analysis, the theory of cyclic extensions Z of an algebraic number field k can be described in terms of cyclic algebras $(Z/k, \lambda, a)$ and collections of local algebras $(Z_p/k_p, \delta, a_p)$, for all prime divisors p of the base field. As a matter of fact, Chevalley's new approach [11] to the classical theory by means of ideal elements may be viewed in the light of our assertion (§8). It is a well-known fact that for arbitrary normal fields K/k with the Galois group Γ, the crossed products $(K/k, \Gamma, F)$, where F denotes a factor set for Γ in K, are the strict analogues of the cyclic algebras. We use this feature of normal fields and their associated algebras to extend the classical (cyclic) theory. In this generalization ideal elements are replaced by "ideal algebras," where an ideal algebra is a collection of local algebras, one for each p-adic extension K_p/k_p (§4).

It might be conjectured from this approach that most of the results of the classical theory may easily be generalized, but this is not the case(2). Entirely new problems, mostly group-theoretical ones, block the path which has been envisaged by various statements of E. Noether [32, 33] on noncommutative methods. One of our final theorems may suffice to illustrate the rather unexpected results of this paper. Our new "class group" of ideal algebras can be represented as $F\mathfrak{A}'/(F'')T\mathfrak{A}'$, where $F\mathfrak{A}'$ denotes the group of factor sets of ideals relatively prime to the different of K/k, where $T\mathfrak{A}'$ is the group of transformation sets of such ideals, and where (F'') are principal ideals generated by "norm residues" (§26). For an abelian (non-cyclic) field this class group is not isomorphic to the Galois group, as in the ordinary theory; it is rather a cyclic group whose order is equal to the least common multiple of the orders of the elements in the Galois group. In other words, for arbitrary abelian fields our theory does not give the classical law of reciprocity. Only a composition which is extraneous to the theory of factor sets yields the usual theory for general abelian extensions.

In the first part of this paper we follow an unpublished investigation of E. Artin (§5). It deals with the theory of p-primary factor sets, that is, factor sets whose elements involve only the divisors of a fixed prime divisor p of k.

Presented to the Society, April 27, 1940; received by the editors August 5, 1940.

(1) The numbers in square brackets refer to the papers in the bibliography.

(2) The first attempts in this direction are due to Tannaka [38].

The results culminate in the following theorem: The least common multiple J of the orders of the elements in Γ is equal to the index $[F\mathfrak{A}' : (F'')T\mathfrak{A}']$. The proof of this theorem is obtained by collecting the results on p-primary factor sets by means of the Tschebotareff density theorem (§§6, 7).

The final chapter of this paper treats of the realization of the above mentioned index relation by a generalized Artin symbol (§29). It is not difficult to generalize the theory of the norm residue symbol (§28). All the formal properties of the old symbol hold true for the new one, except for the usual statements on the values taken by the norm residue symbol for variable argument. The real reason for this discrepancy may be sought in our results on the relation between local algebras and algebras in the large. It is found that in general not every local algebra for $K_\mathfrak{p}/k_\mathfrak{p}$ is the p-component of a suitable crossed product in the large (§§13–16).

The main part of the group-theoretic investigation of the class group is centered around a series of theorems on unit groups (§§17–23). As in the proof of the classical inversion theorem we try to treat the fundamental index $[F\mathfrak{A}' : (F'')T\mathfrak{A}']$ by means of arithmetical reductions of a group-theoretic nature. The class numbers, which are easily canceled off in the old proof, confront us here with a rather involved situation. By means of a group-theoretic invariant, the deficiency index, we finally succeed in disposing of them (§§15, 16). This deficiency index together with the theory of group extensions over unit groups give rise to new invariants for the field K/k. The customary reduction to the Herbrand subgroup can be carried out to a certain extent. However, we find a totally unexpected deviation from the classical case. The principal genus theorem for units and unit groups connects our investigations with Schur's theory of the multiplicator, as was pointed out to us by A. H. Clifford. The group generated by one of the Herbrand units H belonging to an unramified infinite prime divisor can be described in purely group-theoretic terms. If Γ is abelian, the number of group extensions by Γ of such an "abstract unit group" turns out to be exactly the order of the multiplicator of Γ in an algebraically closed field (of characteristic ∞). We wish to emphasize that many of the arithmetic properties of K/k are of strictly group-theoretic nature and can be formulated as such. There are a large number of new problems in group theory which originate in our analysis of normal fields(3). Finally, comparison of the two evaluations for the fundamental index leads to a very complicated relation between the invariants we introduced in the course of our investigation (§25). We discuss this rather mysterious equality in a number of special cases which indicate the relative contributions of the partial indices of $[F\mathfrak{A}' : (F'')T\mathfrak{A}']$. In conclusion we may say that we have explored the potentialities of the concept of factor sets and algebras for arbitrary normal fields.

(3) The statement of these problems in §§17, 21–23 may be read without detailed reference to the remaining sections.

CHAPTER I. FACTOR SETS AND ALGEBRAS

1. **Preliminaries on factor sets.** Let G be a group with an abelian normal subgroup N and a corresponding factor group $\Gamma = G/N$. In each coset σ of G/N choose a representative u_σ, so that the coset σ is Nu_σ. This element u_σ induces in N an automorphism

$$A \leftrightarrow A^\sigma = u_\sigma A u_\sigma^{-1} \qquad (A \text{ in } N). \tag{1}$$

Since N is abelian, any other element v_σ in the coset Nu_σ will induce the same automorphism $A \leftrightarrow A^\sigma$. For σ and τ in Γ, $(A^\tau)^\sigma = A^{\sigma\tau}$ (note the order!).

Conversely, let an abelian group N and a group Γ be given, and let each element σ of Γ be assigned to a definite automorphism $A \leftrightarrow A^\sigma$ of N. We assume that this assignment preserves the products, so that

$$A^{\sigma\tau} = (A^\tau)^\sigma \qquad (\text{all } \sigma, \tau \text{ in } \Gamma). \tag{2}$$

A *group extension* of N by Γ, with the given assignment of automorphisms, is then any group G with N as normal subgroup, and $\Gamma = G/N$ as quotient group, in which each coset σ of Γ induces in N the given automorphism[4] $A \leftrightarrow A^\sigma$.

To represent a group extension explicitly, use a fixed representative u_σ in each coset σ of G/N. The product of two representatives u_σ and u_τ is in the coset of $\sigma\tau$, hence

$$u_\sigma u_\tau = F_{\sigma,\tau} u_{\sigma\tau} \qquad (\text{each } F_{\sigma,\tau} \text{ in } N). \tag{3}$$

The associative law $u_\sigma(u_\tau u_\rho) = (u_\sigma u_\tau)u_\rho$ implies that these constants $F_{\sigma,\tau}$ satisfy the *associativity conditions*

$$F_{\sigma,\tau}F_{\sigma\tau,\rho} = F_{\tau,\rho}^{\sigma} F_{\sigma,\tau\rho} \qquad (\text{all } \sigma, \tau, \rho \text{ in } \Gamma). \tag{4}$$

The equations (1), (3), and (4) determine the group extension G in terms of the subgroup N, quotient group Γ, and constants $F_{\sigma,\tau}$. Any set F of constants $F_{\sigma,\tau}$ in N which satisfy the associativity conditions (4) is called a *factor set*. Every factor set for Γ in N determines a group extension G which consists of elements Au_σ, for A in N, σ in Γ, which are to be multiplied by the rules (1) and (3), or

$$u_\sigma u_\tau = F_{\sigma,\tau} u_{\sigma\tau}, \qquad u_\sigma A = A^\sigma u_\sigma. \tag{5}$$

The *product* of two factor sets F and F' is a third factor set with components $F''_{\sigma,\tau} = F'_{\sigma,\tau}F_{\sigma,\tau}$, for this set F'' clearly satisfies (4). All factors sets F form a group which we denote by $F_\Gamma N$, or simply by[5] FN.

[4] See discussion in Baer [6, 7]; Zassenhaus [43].

[5] We shall adopt this abbreviation if there is no ambiguity about the group Γ. Here and subsequently, starred definitions are given special numbers. They introduce notation important for later arguments.

(*1.1) $F_\Gamma N$ = the group of all factor sets F with components $F_{\sigma,\tau}$ in N.

In a group extension the representatives u_σ of the cosets may be replaced by new representatives $v_\sigma = N_\sigma u_\sigma$, with N_σ in N. This replaces the factor set F by a new set F', given by $F'_{\sigma,\tau} = F_{\sigma,\tau}(N_\sigma N_\tau^\sigma N_{\sigma\tau}^{-1})$. These two factor sets F and F' are called *similar* (notation $F \sim F'$), so

(*1.2) $F \sim F'$ means $F'_{\sigma,\tau} = F_{\sigma,\tau}(N_\sigma N_\tau^\sigma N_{\sigma\tau}^{-1})$,

with suitable elements N_σ of N. Two factor sets determine isomorphic group extensions (with N and Γ fixed) if and only if they are similar.

The added factors $N_\sigma N_\tau^\sigma N_{\sigma\tau}^{-1}$ of (*1.2) themselves form a special sort of factor set known as a *transformation set*. We write

(*1.3) $TN = TN_\sigma$ for a transformation set $N_\sigma N_\tau^\sigma N_{\sigma\tau}^{-1}$, N_σ in N.

If Γ is finite, each such set is derived from a vector

(*1.4) $N_\sigma = \{N_{\sigma_1}, N_{\sigma_2}, \cdots, N_{\sigma_n}\}$,

where each N_{σ_i} is in N and the subscripts σ_i denote the n distinct elements of Γ. The symbol N_σ denotes ambiguously the whole vector or just one component.

LEMMA 1.1. *If Γ has order n, the nth power F^n of any factor set F in $F_\Gamma N$ is a transformation set TN.*

Proof(6). From the given factor set F construct the products

$$C_\sigma = \prod_\tau F_{\sigma,\tau} \qquad \text{(over all } \tau \text{ in } \Gamma\text{)}. \tag{6}$$

The associativity relations (4), multiplied together over all ρ in Γ, become

$$F_{\sigma,\tau}^n C_{\sigma\tau} = C_\tau^\sigma C_\sigma. \tag{7}$$

This states that F^n is the transformation set TC_σ.

2. **Crossed products.** For the basic normal field K of our investigation, we use throughout the following notation:

(*2.1) K, a fixed normal extension of an algebraic number field k,
(*2.2) $n = [K:k]$, the degree of K over k,
(*2.3) A, a nonzero number of K, or (7) the group of all such numbers,
(*2.4) σ, τ, ρ, automorphisms $A \leftrightarrow A^\sigma$ of K/k,
(*2.5) Γ, the Galois group of K/k, composed of all σ,
(*2.6) $F = F_\Gamma A$, a factor set of numbers $F_{\sigma,\tau}$ in A,
(*2.7) (S), a class of normal simple algebras split by K over k.

(6) This proof is apparently originally due to Artin.

(7) As in Hasse's group-stenographic method, a letter may denote ambiguously a group or an arbitrary element of that group.

In connection with this last definition, we recall the properties of algebra classes. Every normal simple algebra B over k has the form of a direct product $B=M\times D$, where D is a division algebra, M a total matric algebra. Two such algebras $B_1=M_1\times D_1$ and $B_2=M_2\times D_2$ are *similar* if and only if D_1 is equivalent to D_2 over k. The set of all algebras similar to B is termed an algebra class (B). The field K is said to *split* an algebra B over k if the algebra B_K formed by extending the coefficient field of B to K is a total matric algebra. If K splits B, it splits every algebra of the class B. The direct *product* of two classes of algebras, taken term-by-term, is a third class $(B_1)\times(B_2)=(B_1\times B_2)$. If both B_1 and B_2 are split by K, so is their product. Therefore the algebra classes (S) of (*2.7) form a group.

Every factor set F in K determines an algebra called a *crossed product* of K by F. It consists of all sums $\sum_\sigma B_\sigma u_\sigma$ formed, with coefficients B_σ in K, from n linearly independent elements u_σ, one for each σ in Γ. These sums are to be added term-by-term and multiplied by the distributive law and the rules (5) of §1 for multiplying the elements u_σ in a group extension. With this convention, the sums $\sum_\sigma B_\sigma u_\sigma$ constitute a normal simple algebra $(K, \Gamma, F)=(K/k, \Gamma, F)$. It has order n^2 over k and contains K as a subfield. Conversely([8]), every normal simple algebra S over k of order n^2 with the subfield K has a crossed product representation (K, Γ, F) for some factor set F. Consequently every algebra class (S) split by K contains a crossed product $S=(K, \Gamma, F)$. Two crossed products belong to the same algebra class if and only if their factor sets are similar. The direct product of two crossed products ([2, pp. 67–73]) is given by the product of the factor sets,

$$(1)\qquad (K/k, \Gamma, F)\times(K/k, \Gamma, F')\sim(K/k, \Gamma, FF').$$

Hence the correspondence $F\to(K, \Gamma, F)$ maps the group F of factor sets homomorphically on the group (S) of algebra classes split by K.

The explicit theory of the algebras S depends on the *valuations* of the base field k: the non-negative real-valued functions $\|a\|$ defined on k with $\|ab\|=\|a\|\|b\|$, $\|a+b\|\leq\|a\|+\|b\|$, $\|a\|=0$ if and only if $a=0$. Two valuations are equivalent if the convergent sequences which they determine are the same; they then determine the same minimal field $k_p\supset k$ *complete* in the convergent sequence topology. By a *prime divisor* (or prime spot) p of k we mean a class of equivalent valuations([9]) of k, so

(*2.8) k_p=the minimal complete field $k_p\supset k$ in the valuation at p.

There are two types of prime divisors. If $a\leftrightarrow a_i$ is an isomorphic mapping of k on the field of complex numbers, so that a_i is a conjugate of a, then the

([8]) See [2, chap. 5], or the parallel discussion in [13, chap. 5]. In our formula (3) of §1 the coefficients $F_{\sigma,\tau}$ appear on the left; in [2] and [13] the corresponding symbols are to the right of u_σ. This causes slight changes in the form of the associativity relations, etc.

([9]) Discussion in [1, pp. 251–305], [13, pp. 93–105], [42, 2d edition, vol. 1].

ordinary absolute value of a_i gives a valuation $\|a\| = |a_i|$ of k. The corresponding prime divisor is called an *infinite* prime divisor $p = p_\infty$, and the associated complete field k_p is either the field of complex numbers or that of real numbers (the latter if the conjugates a_i of all a are real). If p is a prime ideal of the number field k, there is a valuation function $\|a\|_p = \exp[-V_p(a)]$, where

(*2.9) $V_p(a)$ = the exact power to which p divides the ideal (a).

The corresponding prime divisor p is said to be *finite*, and the complete field k_p p-adic. Every prime divisor of k is either finite or infinite in this sense.

Each prime divisor P of the extended field K determines a valuation and hence a prime divisor p of the subfield k; call P a *divisor* of p. Every p has at least one divisor P; the corresponding complete field K_P may be considered as an extension of the original complete field k_p. This *local extension* K_P/k_p is normal, and its Galois group $\Delta(P)$ is a subgroup of the original Galois group Γ of K/k. If P is finite and belongs to a prime ideal P, then $\Delta(P)$ is the Hilbert *decomposition group* of this P, so

(*2.10) $\Delta(P)$ = all δ in the Galois group Γ with $P^\delta = P$.

Conjugate prime ideals P have conjugate groups $\Delta(P)$, so the order of $\Delta(P)$ depends only on the original p,

(*2.11) $m_p = [K_P : k_p] = [\Delta(P) : 1]$.

If P is infinite, K_P is either the real or the complex field. If $K_P = k_p$, $\Delta(P) = 1$, and $P \mid p$ is said to be *unramified*. If K_P is complex, k_p real, $\Delta(P)$ is cyclic of order 2 with a generator $\delta = \delta_P$ and $P \mid p$ is *ramified*.

Consider the corresponding local algebras for any p,

(*2.12) S_p = a normal simple algebra of degree m over k_p, split by K_P/k_p.

A normal simple algebra of degree m over k_p contains an isomorphic map of every possible extension of degree m over k_p (see [9, Lemma 0]), hence S_p contains an unramified cyclic extension W of degree $m = m_p$. Therefore S_p has a cyclic representation $(W/k_p, \sigma, a)$ where σ generates the cyclic group of W/k_p, while a is an element of k_p.

The symbol $(W/k_p, \sigma, a)$ represents the *cyclic algebra* which has over W a basis of m linearly independent elements $1, u_\sigma, u_\sigma^2, \cdots, u_\sigma^{m-1}$, connected by the multiplication table

$$u_\sigma A = A^\sigma u_\sigma, \qquad u_\sigma^m = a; \qquad m = [W : k_p]. \tag{2}$$

For such an algebra one may determine an integer which is invariant (independent of the particular cyclic representation) by the formula(10)

(10) In [22], Hasse uses as invariant the quantity $\rho_p = (1/n)\mu_p$, modulo 1, which is independent of our fixed degree n.

$$\mu(S_p) = \mu(W/k_p, \sigma, a) \equiv (n/m_p)V_p(a) \pmod{n}. \tag{3}$$

Two normal simple algebras over k_p are similar if and only if they have the same invariant μ modulo n.

Another invariant is the index i_p of S_p/k_p, defined as the degree of the division algebra (unique up to isomorphism) which is similar to S_p. This index is known [21, 22] to be the reduced denominator of μ_p/n; that is, it is the smallest integer i_p such that

$$i_p\mu_p \equiv 0 \pmod{n}, \qquad i_p = \text{the index of the algebra } S_p. \tag{4}$$

For an infinite prime divisor p, k_p is either the real or complex field, and the only proper normal division algebra over k_p the algebra of real quaternions Q. The invariant of a normal simple algebra S_p/k_p may then be defined by

$$\begin{aligned} \mu(S_p) &= 0 \qquad \text{if } S_p \text{ is a total matric algebra,} \\ \mu(S_p) &= n/2 \qquad \text{if } S_p \text{ is similar to } Q. \end{aligned} \tag{5}$$

For the direct product $S_p^{(1)} \times S_p^{(2)}$ of two algebras over k_p (p finite or infinite) one always has the formulas

$$\mu(S_p^{(1)} \times S_p^{(2)}) \equiv \mu(S_p^{(1)}) + \mu(S_p^{(2)}) \pmod{n}. \tag{6}$$

An algebra S over k has for each prime divisor p of k a local component $(S)_p = S_p = S \times k_p$, obtained by extending the coefficient field to the complete field k_p. This algebra S_p is normal simple over k_p. Hence S has a set of invariants $\mu_p(S) = \mu(S_p)$, one for each p. Conversely, it is known that S is determined up to similarity by the set of its invariants [2, 21, 22]. If S_p is a total matric algebra, $S_p \sim k_p$, then S is said to be unramified at p, otherwise ramified. It is known that S is ramified at only a finite number of prime divisors ([2, pp. 148–149]).

For a crossed product the local algebras can be explicitly represented ([21, 22]) as

$$S_p = (K/k, \Gamma, F)_p \sim (K_P/k_p, \Delta(P), F \cap \Delta) \tag{7}$$

where $F \cap \Delta$ denotes the part of the factor set F applying to the group $\Delta(P) = \Delta$; that is, $F \cap \Delta$ is the factor set with terms $F_{\xi,\eta}$ for ξ, η in Δ. Here Δ may be the decomposition group for any one of the prime divisors P of p. If P is finite and $P \mid p$ unramified, the invariant μ_p may be explicitly calculated from the factor set F. For, under these circumstances the local extension K_P/k_p is cyclic, and its group Δ has as generator

(*2.13) $\delta = \delta_P$ = the Frobenius automorphism $[(K/k)/P]$ which is characterized by the property[11]

[11] See [20, Part Ia, p. 71]; [23, pp. 36, 38].

(8) $A^{\delta} \equiv A^{Np} \pmod{P}$, Np = absolute norm of p, $V_P(A) \geqq 0$.

Because K_P/k_p is cyclic, the crossed product $(K_P/k_p, \Delta(P), F \cap \Delta)$ can be written as a cyclic algebra $(K_P/k_p, \delta_P, a)$, while the generator u_δ for this algebra may be chosen as the corresponding basis element u_δ of the given crossed product. By (2) the multiplication constant a is then u_δ^m, where m is the p-degree $m = m_p$ of (*2.11). However

$$a = u_\delta^m = (u_\delta u_\delta)u_\delta^{m-2} = F_{\delta,\delta}(u_\delta^2 u_\delta)u_\delta^{m-3} = \cdots = F_{\delta,\delta}F_{\delta^2,\delta}\cdots F_{\delta^{m-1},\delta}u_{\delta^m}.$$

But $u_{\delta^m} = u_1$, and $u_1 u_\delta = F_{1,\delta}u_\delta$, so $u_1 = F_{1,\delta}$. Consequently

$$a = F_{1,\delta}F_{\delta,\delta}\cdots F_{\delta^{m-1},\delta} = \prod_{i=0}^{m-1} F_{\delta^i,\delta} \qquad (\delta = \delta_P). \tag{9}$$

According to the definition (3) of the invariant μ_p, we have

$$\mu_p(K, \Gamma, F) = (n/m)V_P\left[\prod_{i=0}^{m-1} F_{\delta^i,\delta}\right], \tag{10}$$

where $\delta = \delta_P$ is the Frobenius-Artin automorphism of order $m = m_p$.

3. **The module** M. As in the cyclic case, one must restrict the numbers under consideration to those relatively prime to some divisor

$$M = P_1^{h_1}P_2^{h_2}\cdots P_s^{h_s}, \tag{1}$$

where each P_i denotes a finite or infinite prime divisor of K, while the h_i are positive integers. The requirement "A relatively prime to M" means that A is relatively prime to each factor $P_i^{h_i}$ in M. If P_i is a finite prime divisor, this statement has its customary meaning, while "A is relatively prime to an infinite P" means simply "$A \neq 0$." It is convenient to say that a prime divisor p of k is *involved* in M if some factor P of p is present in M (with a positive exponent).

For M in the subsequent developments we use any module which satisfies the conditions

(i) any p ramified in K/k is involved in M,
(ii) if p is involved in M, every $P \mid p$ is present in M.

For example, M might be the product of all P's ramified in K/k. Alternatively, M might be the factor-set conductor of K/k which is defined in §26.

A prime "$'$" will be used to denote elements or groups of elements relatively prime to the module M.

(*3.1) A' = the group of numbers $A \neq 0$ of K relatively prime to M.

(*3.2) F' = the group of all factor sets of numbers $F_{\sigma,\tau}$ relatively prime to M.

(*3.3) $S'=S'(K)=$the group of all normal simple algebras S relatively prime to M.

Here we call an algebra S "relatively prime to M" if none of the ramification divisors of S are involved in M; that is, if p involved in M implies $S\times k_p\sim 1$ [12].

4. **Ideal algebras.** By an *ideal algebra* $\mathfrak{S}$ over the field k we simply mean a specification of normal simple algebras S_p over k_p, one for each prime divisor p of k, such that $S_p\sim k_p$ except for a finite number of prime divisors. The algebra S_p is the *component* $(\mathfrak{S})_p=\mathfrak{S}_p$ of $\mathfrak{S}$. We agree to call two such ideal algebras $\mathfrak{S}_1$ and $\mathfrak{S}_2$ equal if and only if their components $(\mathfrak{S}_1)_p\sim(\mathfrak{S}_2)_p$ are similar for every p. For each such p we have the corresponding p-index $i_p(\mathfrak{S})=i(S_p)$ of the local algebra S_p. We say that $\mathfrak{S}$ is *split* by the given extension K/k if, for each p, $i_p(\mathfrak{S})$ is a divisor of the p-degree m_p of K/k. In case $\mathfrak{S}$ has the components $\mathfrak{S}_p=S_p$ of an actual algebra S, this condition is simply the usual condition that the algebra S be split by K [21, 22, 2]. Ideal algebras form an abelian group under the operation of the direct product

$$(\mathfrak{S}_1\times\mathfrak{S}_2)_p=S_{1p}\times S_{2p}. \tag{1}$$

Under this operation the set of all ideal algebras split by K constitutes a subgroup

(*4.1) $\mathfrak{S}(K)=\mathfrak{S}=$the group of all $\mathfrak{S}$'s split by K,

(*4.2) $\mathfrak{S}'(K)=\mathfrak{S}'=$the group of all $\mathfrak{S}$'s relatively prime to M and split by K.

Again, $\mathfrak{S}$ relatively prime to M means $S_p\sim k_p$ for all p involved in M. The group $S(K)=S$ of actual algebras forms, in natural fashion, a subgroup of $\mathfrak{S}$. We propose to compute the index

(*4.3) $J=[\mathfrak{S}':S']$.

An ideal algebra $\mathfrak{S}$ is determined uniquely by giving the set of its local invariants $\mu_p=\mu_p(\mathfrak{S})=\mu(S_p)$, for each prime divisor p. Thus an ideal algebra $\mathfrak{S}$ split by K is completely specified by a list of integers μ_p subject to the conditions

$$m_p\mu_p\equiv 0 \pmod n \qquad \text{for all } p, \tag{2}$$

$$\mu_p\equiv 0 \pmod n \qquad \text{for all but a finite number of prime divisors } p. \tag{3}$$

THEOREM 4.1. *The index J defined in* (*4.3) *is $J(\Gamma)$, the least common multiple of the orders of the elements of the Galois group Γ.*

Proof. The only relation between the invariants of a crossed product is $[2, 13, 22]\sum_p\rho_p(S)\equiv 0 \pmod 1$. In terms of the invariants μ_p, this becomes

[12] Remember that $(K/k, \Gamma, F')_p\sim 1$ need not hold for p involved in M.

$\sum_p \mu_p(S) \equiv 0 \pmod{n}$. For a given ideal algebra the quantity

$$d(\mathfrak{S}) \equiv \sum_p \mu_p(\mathfrak{S}) \pmod{n} \tag{4}$$

measures the divergence of $\mathfrak{S}$ from "actuality." More explicitly, the map $\mathfrak{S} \to d(\mathfrak{S})$ carries the group $\mathfrak{S}$ homomorphically (see §2, (6)) into an additive group of integers d mod n, in such a fashion that the subgroup carried into 0 mod n is the group of actual algebras. By the restriction (2), each $d(\mathfrak{S}')$ satisfies the condition $J(\Gamma)d(\mathfrak{S}') \equiv 0 \pmod{n}$. Therefore $\mathfrak{S} \to d(\mathfrak{S})$ maps the quotient group $\mathfrak{S}'/S'$ isomorphically on the cyclic group of integers of order $J(\Gamma)$, generated by $n/J(\Gamma) \pmod{n}$, provided there exists an ideal algebra $\mathfrak{S}'$ with $d(\mathfrak{S}') = n/J(\Gamma)$. It remains to construct this $\mathfrak{S}$. Recall that $J(\Gamma)$ is the least common multiple of the orders $O_1, \cdots, O_t$ of certain elements $\sigma_1, \cdots, \sigma_t$ of the Galois group. By partial fractions, we may express $1/J(\Gamma)$ as

$$1/J(\Gamma) = a_1/O_1 + \cdots + a_t/O_t \qquad (a_i \text{ integers}).$$

The Tschebotareff density theorem (see [20, Part II, §24]) asserts that for a given σ_i there are (infinitely many) sets of distinct prime divisors p_i of k with factors P_i in K such that the Frobenius automorphism of P_i is σ_i. One may assume each p_i to be relatively prime to M. By definition, the Frobenius automorphism generates the decomposition group $\Delta(P_i)$, so that the order m_i of this group is the given order O_i of σ_i. There is thus an ideal algebra $\mathfrak{S}$ relatively prime to M and split by K, with invariants

$$\mu_{p_i} \equiv na_i/O_i \pmod{n}, \qquad i = 1, 2, \cdots, t,$$
$$\mu_q \equiv 0 \pmod{n}, \qquad (q \neq p_1, p_2, \cdots, p_t).$$

For this algebra $\mathfrak{S}$ we get $d(\mathfrak{S}) = n/J(\Gamma)$, as desired for Theorem 4.1.

In this proof we may drop the requirement that the ideal algebras $\mathfrak{S}$ be relatively prime to M. The result is analogous.

THEOREM 4.2. *The index* $[\mathfrak{S}:S]$ *of the group of actual algebras in the group of ideal algebras split by* K *equals* $J_0(\Gamma)$, *the least common multiple of the orders* m_p *of the decomposition groups of the primes* p *in* K/k.

5. **Artin's character method.** An ideal algebra can be given by a factor set of ideals in K, as we now show by computing the "*invariants*" μ_p of such a factor set. We use the following notation, in which each letter may also denote the group of all objects so labelled, while a prime "$'$" will denote the restriction "relatively prime to M":

(*5.1) $\mathfrak{A}$ = an ideal of K (possibly a fractional ideal),
(*5.2) $\mathfrak{A}_\sigma$ = a vector of n ideals, one for each σ in Γ,
(*5.3) $T\mathfrak{A}$ = a transformation set $\mathfrak{A}_\sigma \mathfrak{A}_\tau^{\sigma} \mathfrak{A}_{\sigma\tau}^{-1}$ derived from $\mathfrak{A}_\sigma$,
(*5.4) $\mathfrak{F}$ = a factor set $\mathfrak{F}_{\sigma,\tau}$ of ideals of K, with σ, τ in Γ.

For each such set we introduce a function $h_P(\sigma)$, determined by σ in Γ and a prime divisor P of K,

(*5.5) $h_P(\sigma) = h_P(\sigma, \mathfrak{F}) \equiv \sum_\tau V_P(\mathfrak{F}_{\sigma,\tau}) \pmod{n}$,

where V_P is the P-adic valuation (see (*2.9)). If P is unramified in K/k, we propose as an invariant of $\mathfrak{F}$ the integer

(*5.6) $\mu_P = \mu_P(\mathfrak{F}) \equiv h_P(\delta, \mathfrak{F}) \pmod{n}$,

where $\delta = \delta_P$ is the Frobenius automorphism (see (*2.13)).

If p is the prime ideal of k divisible by P, each ideal in a factor set $\mathfrak{F}$ may be factored as

$$\mathfrak{F}_{\sigma,\tau} = \mathfrak{F}_{\sigma,\tau}^{(p)} \cdot \mathfrak{B}_{\sigma,\tau}, \qquad (\mathfrak{B}_{\sigma,\tau}, p) = 1 \tag{1}$$

where the first factor $\mathfrak{F}_{\sigma,\tau}^{(p)}$ involves only the prime factors $P_1, P_2, \cdots$ of p. These ideals $\mathfrak{F}_{\sigma,\tau}^{(p)}$ constitute by themselves a factor set which is p-primary in the sense of the

DEFINITION. *An ideal factor set $\mathfrak{F}$ is called p-primary if all prime ideal factors of any ideal $\mathfrak{F}_{\sigma,\tau}$ of the set are prime ideal factors P of the given p.*

The factor set $\mathfrak{F}$ and its p-primary component $\mathfrak{F}^{(p)}$ clearly determine the same function $h_p(\sigma)$[12].

LEMMA 1. *For factor sets of ideals $\mathfrak{F}_1$ and $\mathfrak{F}_2$, $\mathfrak{F}_1 \sim \mathfrak{F}_2$ implies $h_P(\delta, \mathfrak{F}_1) \equiv h_P(\delta, \mathfrak{F}_2) \pmod{n}$ for all δ in $\Delta(P)$.*

Proof. Multiplication of factor sets is represented by addition of the corresponding functions h, for the definition shows that

$$h_P(\sigma, \mathfrak{F}_1\mathfrak{F}_2) \equiv h_P(\sigma, \mathfrak{F}_1) + h_P(\sigma, \mathfrak{F}_2) \pmod{n}. \tag{2}$$

But $\mathfrak{F}_1 \sim \mathfrak{F}_2$ means that $\mathfrak{F}_1 = \mathfrak{F}_2\mathfrak{F}$, where $\mathfrak{F}$ is a transformation set. Hence we need only prove $h_P(\delta, \mathfrak{F}) \equiv 0 \pmod{n}$, whenever $\mathfrak{F}$ is a transformation set $T\mathfrak{A}$. But

$$\prod_\tau [\mathfrak{A}_\tau \mathfrak{A}_{\tau'}^{\sigma} \mathfrak{A}_{\sigma\tau}^{-1}] = \mathfrak{A}_\sigma^{\sigma} \left(\prod_\tau \mathfrak{A}_\tau \right)^{\sigma} \cdot \left(\prod_{\tau'} \mathfrak{A}_{\tau'} \right)^{-1} = \mathfrak{A}_\sigma^{\sigma} \left(\prod_\tau \mathfrak{A}_\tau \right)^{\sigma-1}. \tag{3}$$

For σ in $\Delta(P)$, $V_P(\prod_\tau \mathfrak{A}_\tau) = V_P(\prod_\tau \mathfrak{A}_\tau)^\sigma$, since σ leaves P fixed. Hence

$$h_P(\delta, \mathfrak{A}_\sigma \mathfrak{A}_{\tau'}^{\sigma} \mathfrak{A}_{\sigma\tau}^{-1}) = nV_P(\mathfrak{A}_\sigma) \equiv 0 \pmod{n}.$$

---

[12] The function h and its properties are due to unpublished work of E. Artin. The method used will apply to factor sets $\mathfrak{F}$ in any $F_\Gamma(\mathfrak{A})$ for which the abstract group $\mathfrak{A}$ with operators has a suitable structure (generated by free generators $P, Q, \cdots$, suitably permuted by Γ); see the statement of Theorem 5 in [17].

LEMMA 2. *The function h corresponding to any factor set $\mathfrak{F}$ is a character of $\Delta(P)$ mod n; which is to say that*

$$h_P(\zeta) + h_P(\eta) \equiv h_P(\zeta\eta) \pmod{n} \qquad (\zeta, \eta \text{ in } \Delta(P)). \tag{4}$$

Proof. As in §1, (7) and (8), we introduce a vector $\mathfrak{C}_\sigma$ for each $\mathfrak{F}$, with

$$\mathfrak{C}_\sigma = \prod_\tau \mathfrak{F}_{\sigma,\tau}, \qquad \mathfrak{F}^n_{\sigma,\tau} = \mathfrak{C}_\sigma \mathfrak{C}_\tau^\sigma \mathfrak{C}_{\sigma\tau}^{-1}. \tag{5}$$

The function $h_P(\sigma)$ is then given by

$$h_P(\sigma) \equiv V_P(\mathfrak{C}_\sigma) \pmod{n}. \tag{6}$$

By (5), the transformation set $T\mathfrak{C}$ consists of nth powers, which means that

$$V_P(\mathfrak{C}_\sigma) + V_P(\mathfrak{C}_\tau^\sigma) \equiv V_P(\mathfrak{C}_{\sigma\tau}) \pmod{n}. \tag{7}$$

But $V_P(\mathfrak{C}_\tau^\sigma)$, the exponent to which the prime P appears in $\mathfrak{C}_\tau^\sigma$, is simply the exponent of the prime $P^{\sigma^{-1}}$ in $\mathfrak{C}_\tau$. If we set $\sigma^{-1} = \rho$, (7) becomes

$$V_{P^\rho}(\mathfrak{C}_\tau) \equiv h_P(\rho^{-1}\tau) - h_P(\rho^{-1}) \pmod{n}. \tag{8}$$

If in (7) we let σ be an element δ^{-1} in the decomposition group $\Delta(P)$, the result, expressed in terms of h, is

$$h_P(\delta) + h_P(\tau) \equiv h_P(\delta\tau) \pmod{n}, \qquad (\delta \text{ in } \Delta(P), \tau \text{ in } \Gamma). \tag{9}$$

This includes the desired conclusion (4) as a special case. By a similar device we may prove a partial converse to Lemma 1.

LEMMA 3. *If $h_P(\delta, \mathfrak{F}^{(p)}) \equiv 0 \pmod{n}$, for all δ in $\Delta(P)$, then the primary factor set $\mathfrak{F}^{(p)}$ is a transformation set.*

Proof. By (9), $h_P(\delta\tau) \equiv h_P(\tau) \pmod{n}$ then depends only on the coset of τ modulo $\Delta(P)$. There is therefore an ideal $\mathfrak{L}$ divisible by P^ρ exactly to the power $h_P(\rho^{-1})$, for each factor P^ρ of p, or

$$V_{P^\rho}(\mathfrak{L}) = h_P(\rho^{-1}, \mathfrak{F}).$$

Using (8), one may then show that $V_{P^\rho}(\mathfrak{L}^{1-\tau}\mathfrak{C}_\tau) \equiv 0 \pmod{n}$ for every ρ, which is to say that $\mathfrak{L}^{1-\tau}\mathfrak{C}_\tau$ is the nth power of some vector of ideals $\mathfrak{B}_\sigma$,

$$\mathfrak{L}^{1-\tau}\mathfrak{C}_\tau = \mathfrak{B}_\tau^n \qquad (\text{all } \tau \text{ in } \Gamma).$$

Substitution of this value of $\mathfrak{C}_\tau$ in the second equation of (5) yields $\mathfrak{F}^n = T\mathfrak{B}^n$. Since the nth roots can be extracted uniquely (if at all) in the group of ideals, the last equation means that $\mathfrak{F} = T\mathfrak{B} \sim 1$, as asserted.

LEMMA 4. *Every character $h(\delta)$ of the decomposition group $\Delta(P)$ is the character $h(\delta) = h_P(\delta, \mathfrak{F})$ of some p-primary factor set $\mathfrak{F}$ of ideals.*

Proof. In Γ select for each coset $\sigma\Delta(P)$ a representative σ^*, taking care to select the identity as the representative of $\Delta(P)$ itself. A product $\sigma\sigma^{*-1}$ is then always in $\Delta(P)$. From the given character $h(\delta)$ of $\Delta(P)$ we define an extended function $h^*(\sigma)$ as

$$h^*(\sigma) = h(\sigma\sigma^{*-1}). \tag{10}$$

By the assumption (4) on h one then computes that h^* satisfies (9). This implies that the expression $h^*(\sigma\tau) - h^*(\tau)$ is unaltered, modulo n, by replacement of σ by another element $\delta\sigma$ in the same coset. There must therefore be a vector of ideals $\mathfrak{C}_\tau$ with

$$\begin{aligned} V_P(\mathfrak{C}_\tau) &= h^*(\tau) - h^*(1), \\ V_{P^\rho}(\mathfrak{C}_\tau) &\equiv h^*(\rho^{-1}\tau) - h^*(\rho^{-1}) \pmod{n}, \end{aligned} \tag{11}$$

in analogy with (8). From this definition one computes that $V_{P^\rho}(\mathfrak{C}_\sigma \mathfrak{C}_\tau^\sigma \mathfrak{C}_{\sigma\tau}^{-1}) \equiv 0 \pmod{n}$, so that this transformation set must be an nth power $T\mathfrak{C} = \mathfrak{F}^n$. Here $\mathfrak{F}$ must be a factor set because it is the (unique) nth root of the known factor set $T\mathfrak{C}$. Furthermore this p-primary factor set $\mathfrak{F}$ has the given function $h(\delta)$ as character, for one computes by (11) that $h_P(\delta) = V_P(\mathfrak{C}_\delta) = h(\delta) - h(1) \equiv h(\delta) \pmod{n}$.

These results may be summarized by

THEOREM 5.1. *If the prime ideal factor P of p in K/k has the decomposition group $\Delta(P)$ with the commutator subgroup $\Delta(P)'$, then the number of classes of p-primary factor sets of ideals in K is equal to the index* $[\Delta(P):\Delta(P)']$.

Proof. The preceding lemmas show that the correspondence $\mathfrak{F} \to h_P(\delta, \mathfrak{F})$ maps the group of classes of p-primary factor sets isomorphically on the group of characters $h(\delta)$ mod n of $\Delta(P)$. Hence we need only count the number of such characters. Each character maps the commutator subgroup $\Delta(P)'$ onto zero, mod n. Furthermore, each character of $\Delta(P)$ is induced by a character of the abelian quotient group $\Delta(P)/\Delta(P)'$. Conversely, this abelian group has a number of characters equal to its order, which is the index $[\Delta(P):\Delta(P)']$ of the conclusion.

COROLLARY. *If $P \mid p$ is unramified, the number of classes of p-primary factor sets is the p-degree m_p of K_P over k_p.*

Proof. $P \mid p$ unramified makes $\Delta(P)$ cyclic. The character $h_P(\delta)$ is essentially independent of the choice of the prime ideal factor P of p. If P^ρ is any other factor, the decomposition group $\Delta(P^\rho)$ of P^ρ is obtained from $\Delta(P)$ by the isomorphism $\delta \to \rho\delta\rho^{-1}$. If δ is the Frobenius automorphism of P, then $\rho\delta\rho^{-1}$ is the Frobenius automorphism of P^ρ.

LEMMA 5. *For any prime factor $Q = P^\rho$, the corresponding character is*

$$h_Q(\rho\delta\rho^{-1}, \mathfrak{F}) = h_P(\delta) \qquad (\delta \text{ in } \Delta(P)). \tag{12}$$

If $P|p$ is unramified, P and Q determine the same invariant $\mu_P(\mathfrak{F}) = \mu_Q(\mathfrak{F})$.

For by (6), (8), and (9),

$$h_Q(\rho\delta\rho^{-1}) = V_Q(\mathfrak{C}_{\rho\delta\rho^{-1}}) = V_P(\mathfrak{C}_{\rho\delta\rho^{-1}}^{p^{-1}})$$
$$= V_P(\mathfrak{C}_{\delta\rho^{-1}}) - V_P(\mathfrak{C}_{\rho^{-1}}) \equiv h_P(\delta\rho^{-1}) - h_P(\rho^{-1}) \equiv h_P(\delta) \pmod{n}.$$

It is useful to express the character h in several different forms.

LEMMA 6. *The character of any factor set $\mathfrak{F}$ is given by the expressions*

$$h_P(\delta, \mathfrak{F}) \equiv (n/m) \sum_{\eta} V_P(\mathfrak{F}_{\delta,\eta}) \qquad (\text{summed over all } \eta \text{ in } \Delta), \tag{13}$$
$$\equiv (n/m) \sum_{\eta} V_P(\mathfrak{F}_{\eta,\delta}) \qquad (\text{summed over all } \eta \text{ in } \Delta),$$

where $m = m_P$ is the order of the decomposition group $\Delta = \Delta(P)$. If ζ lies in a subgroup $\Lambda \subset \Delta$ of order r, then

$$h_P(\zeta, \mathfrak{F}) \equiv (n/r) \sum_{\eta} V_P(\mathfrak{F}_{\eta,\zeta}) \qquad (\text{summed over all } \eta \text{ in } \Delta). \tag{14}$$

Proof. In the second half of (5), set $\sigma = \delta$, $\tau = \eta$, take the order V_P and sum over η (i.e., over $\delta\eta$). There results

$$n \sum_{\eta} V_P(\mathfrak{F}_{\delta,\eta}) = mV_P(\mathfrak{C}_\delta).$$

By (6), we have on the right $mh_P(\delta)$, hence the first result of (13). For convenience, we represent the exact exponent of P in the decomposition of $\mathfrak{F}_{\zeta,\eta}$ by

$$e_{\zeta,\eta} = V_P(\mathfrak{F}_{\zeta,\eta}).$$

The second half of (13) reduces to proving that the "right" and "left" sum functions

$$R(\zeta) = \sum_{\eta} e_{\zeta,\eta}, \qquad L(\zeta) = \sum_{\eta} e_{\eta,\zeta}$$

are identical. But the associativity relations for $\mathfrak{F}$ yield for the exponents $e_{\zeta,\eta}$ the analogous relations

$$e_{\zeta,\eta} + e_{\zeta\eta,\xi} = e_{\eta,\xi} + e_{\zeta,\eta\xi} \qquad (\zeta, \eta, \xi \text{ in } \Delta(P)). \tag{15}$$

If this equation is added over ζ and then over ξ, one finds

$$L(\eta) + L(\xi) = me_{\eta,\xi} + L(\eta\xi), \quad me_{\zeta,\eta} + R(\zeta\eta) = R(\eta) + R(\zeta).$$

If we solve these equations for the common value $me_{\zeta,\eta}$, we find

$$R(\zeta) + R(\eta) - R(\zeta\eta) = L(\zeta) + L(\eta) - L(\zeta\eta).$$

In this equation set $\eta=\delta^i$, $\zeta=\delta$, and add for $i=0, 1, \cdots, m-1$. The result is $mR(\delta)=mL(\delta)$, and hence $R(\delta)=L(\delta)$, as desired for (13).

The relation (14) describes in effect the behaviour of a character when applied to a subfield (corresponding to the subgroup Λ) of K. We set

$$L_\Lambda(\zeta) = \sum_\eta e_{\eta,\zeta} \qquad \text{(summed over all } \eta \text{ in } \Lambda).$$

The associativity relations (15), summed over all ζ in Λ, give

$$L_\Lambda(\eta) + L_\Lambda(\xi) = re_{\eta,\xi} + L_\Lambda(\eta\xi) \qquad (\eta, \xi \text{ in } \Lambda).$$

Combination with the similar equation for L itself yields

$$(m/r)\left[L_\Lambda(\eta) + L_\Lambda(\xi) - L_\Lambda(\eta\xi)\right] = L(\eta) + L(\xi) - L(\eta\xi).$$

As in the previous case this entails $(m/r)L_\Lambda(\eta)=L(\eta)$. This gives (14), for the functions L can be written as

$$L_\Lambda(\zeta) = \sum_\eta V_P(\mathfrak{F}_{\eta,\zeta}) \qquad (\eta \text{ in } \Lambda),$$

$$L(\zeta) = \sum_\eta V_P(\mathfrak{F}_{\eta,\zeta}) \qquad (\eta \text{ in } \Delta(P)).$$

The invariant $\mu_P(\mathfrak{F})$ as defined in (*5.2) is the same for all prime factors Q of p by Lemma 5, for it is known (see [20, Part II, p. 51]) that the Frobenius automorphism for $Q=P^\rho$ is obtained from that of P as $\delta_Q=\rho\delta_P\rho^{-1}$. We write $\mu_p(\mathfrak{F})=\mu_P(\mathfrak{F})=\mu_Q(\mathfrak{F})$ for this common value. These invariants of $\mathfrak{F}$ form a complete set for ideals relatively prime to the module M.

THEOREM 5.2. *Two factor sets $\mathfrak{F}_1'$, $\mathfrak{F}_2'$ satisfy $\mathfrak{F}_1'\sim\mathfrak{F}_2'$ if and only if $\mu_p(\mathfrak{F}_1')\equiv\mu_p(\mathfrak{F}_2')$ (mod n) for every p relatively prime to M.*

Proof. According to Lemma 1 we need only show that a factor set $\mathfrak{F}$ with its invariants all zero is a transformation set. As in (1), consider some p-primary component $\mathfrak{F}^{(p)}$ of the given factor set. Since the ideals are all relatively prime to the module M, we need consider here only prime divisors p which are unramified in K/k. For such an unramified p the decomposition group $\Delta(P)$ is cyclic with generator δ_P, and the invariant $\mu_p=h_P(\delta_P)$ determines the whole character $h(\delta_P)$. Hence $\mu_p\equiv 0$ (mod n) implies $h_P(\delta_P)\equiv 0$ (mod n), which in turn makes $\mathfrak{F}^{(p)}\sim 1$, as in Lemma 3. Since this holds for any p-component, $\mathfrak{F}$ itself is ~ 1. In similar fashion we have

THEOREM 5.3. *Let integers μ_p be given for every p relatively prime to M such that, for every p,*

$$m_p\mu_p \equiv 0 \text{ (mod } n),$$

and such that $\mu_p\equiv 0$ (mod n) except for a finite number of prime divisors p. Then there exists a factor set $\mathfrak{F}'$ of ideals relatively prime to M with the invariants μ_p.

The computation of μ_p may be summarized (see [13, 14]) by

LEMMA 7. *If $\delta = \delta_P$ is the Frobenius automorphism of a factor P of p, then*

$$\mu_p(\mathfrak{F}) = (n/m)V_P\left[\prod_{i=0}^{m-1} \mathfrak{F}^{\delta^i,\delta}\right], \qquad m = m_p. \tag{16}$$

Furthermore, for the product of two factor sets

$$\mu_p(\mathfrak{F}_1\mathfrak{F}_2) \equiv \mu_p(\mathfrak{F}_1) + \mu_p(\mathfrak{F}_2) \pmod{n}. \tag{17}$$

If $\eta = \delta^t$ generates in $\Delta(P)$ a subgroup of order r, then

$$t\mu_p(\mathfrak{F}) \equiv (n/r)V_P\left[\prod_{i=0}^{r-1} \mathfrak{F}^{\eta^i,\eta}\right] \pmod{n}. \tag{18}$$

Note that formula (16) is just like the formula (10) in §2 for the invariant μ_p of an actual algebra. Hence

THEOREM 5.4. *Any factor set FA of numbers determines a factor set $\mathfrak{F} = (FA)$ of principal ideals, such that for unramified p the invariants μ_p of the ideal factor set $\mathfrak{F}$ and the algebra (K, Γ, FA) coincide:*

$$\mu_p((FA)) \equiv \mu_p(K, \Gamma, FA) \pmod{n}. \tag{19}$$

It is instructive to observe that the analysis of p-primary factor sets may be reduced to the factor sets $e_{\zeta,\eta}$ of integers with the associativity conditions (15). Such a factor set determines a group extension of the additive group E of rational integers by the group $\Delta(P)$, under the assumption that each ζ in $\Delta(P)$ induces in E the identity automorphism $e^\zeta = e$, e in E. One may prove

THEOREM 5.5. *If P is a prime ideal factor of p, and if $\mathfrak{F}_{\sigma,\tau} = P'^{\sigma,\tau}\mathfrak{B}_{\sigma,\tau}$, with $(\mathfrak{B}_{\sigma,\tau}, P) = 1$, is a p-primary factor set $\mathfrak{F}$ of ideals, then the exponents $f_{\zeta,\eta}$ for ζ, η in the decomposition group form an additive factor set $FE = f_{\zeta,\eta}$ of integers. The correspondence $\mathfrak{F} \to FE$ maps the classes of similar factor sets $\mathfrak{F}$ isomorphically on the classes of similar factor sets of integers.*

COROLLARY. *The number of group extensions of the group of p-primary ideals by the Galois group Γ is the same as the number of group extensions of the additive group of integers by the decomposition group $\Delta(P)$.*

Here a "p-primary ideal" is one whose prime ideal factors are all factors of the given prime ideal p of k.

6. **Factor sets for ideal algebras.** Following E. Noether, we consider the factor sets given by

(*6.1) F'' = factor sets FA composed of numbers $F_{\sigma,\tau}$ relatively prime to M, and such that the algebra $S = (K, \Gamma, FA)$ is relatively prime to M.

As before, "S relatively prime to M" means $S_p \sim k_p$ for every p involved in M.

Such factor sets F'' determine certain factor sets of principal ideals $(F_{\sigma,\tau})$ which are said to constitute the *principal class*[14] of ideal factor sets.

(*6.2) (F'') = the group of factor sets $(F_{\sigma,\tau})$ of principal ideals generated by[15] factor sets $F_{\sigma,\tau}$ of type F''.

This group is contained in the group $\mathfrak{F}'$ of all factor sets of ideals relatively prime to[16] M.

The representation of an ideal algebra $\mathfrak{S}'$ by an ideal factor set in $\mathfrak{F}' = F\mathfrak{A}'$ may be stated as

THEOREM 6.1. *The index J (the number of ideal algebras relatively prime to M modulo the actual algebras relatively prime to M) is given by*

$$J = [\mathfrak{S}' : S'] = [F\mathfrak{A}' : T\mathfrak{A}' \cdot (F'')], \tag{1}$$

where $T\mathfrak{A}' \cdot (F'')$ denotes the join of the subgroup $T\mathfrak{A}'$ of transformation sets and the subgroup (F'') described above.

Proof. Each factor set $\mathfrak{F}' = F\mathfrak{A}'$ of ideals determines a set of invariants $\mu_p(\mathfrak{F}')$ which are the invariants of some ideal algebra split by K. The correspondence

$$\mathfrak{F}' \to \mathfrak{S}' \quad \text{if, for all } p, \quad \mu_p(\mathfrak{F}') \equiv \mu_p(\mathfrak{S}') \pmod{n} \tag{2}$$

maps the group of factor sets $\mathfrak{F}'$ homomorphically into the group of ideal algebras $\mathfrak{S}'$ relatively prime to M, for multiplication of factor sets corresponds to multiplication of algebras (§5, (17), §2, (6)). For each $\mathfrak{S}'$ there is a corresponding set of invariants μ_p, for which there must be, by Theorem 5.3, a corresponding ideal factor set $\mathfrak{F}'$. Hence (2) is a homomorphism. To prove the identity (1), it remains only to investigate the subgroup of $\mathfrak{F}$ carried by (2) into the group of actual algebras.

The subgroup $T\mathfrak{A}'$ of transformation sets is clearly carried by (2) into a certain group of actual algebras, for the invariants of a transformation set are all zero (mod n), as stated in Theorem 5.2, so the corresponding algebra is similar to k. On the other hand, the principal ideal factor set (F'') is mapped by (1) on the actual algebra (K, Γ, F''), according to Theorem 5.4. It remains to show that every factor set mapped into S' lies in the product $T\mathfrak{A}' \cdot (F'')$. If $F \to S'$, then the algebra S' has a crossed product representation (K, Γ, F) in which the factor set F may be so chosen that its components are relatively prime to[17] M. Thus F is a factor set F''. According to the correspondence

---

(14) This principal class is narrower than that defined by Noether in [33], to the exact extent of the requirement that each $F_{\sigma,\tau}$ be relatively prime to M.

(15) In this definition, it is essential not only that $(F_{\sigma,\tau})$ be an ideal factor set, but also that the numbers $F_{\sigma,\tau}$ be themselves a factor set.

(16) Note that the infinite prime divisors in M do not impose conditions on ideals.

(17) Proof is given below; see Theorem 6.2.

(2), $\mathfrak{F}$ and S' have the same invariants. Hence, by Theorem 5.4, $\mathfrak{F}$ and (F'') have the same invariants. Therefore their quotient is a transformation set (see Theorem 5.2). This proves (1) completely.

In the above proof we have assumed part of the following characterization of factor sets relatively prime to M.

THEOREM 6.2. *Let S be a normal simple algebra over k with maximal subfield K. Then S has a crossed product representation $S=(K, \Gamma, F')$ with a factor set F' of numbers relatively prime to the module M if and only if, for every prime ideal $\mathfrak{p}$ involved in M, the $\mathfrak{p}$-adic component $S_\mathfrak{p}$ of S has a representation*

$$S_\mathfrak{p} \sim (K_P/k_\mathfrak{p}, \Delta(P), FE_P) \tag{3}$$

in which FE_P is a factor set for the group $\Delta(P)$ in the group E_P of units of the field K_P[18].

Proof. If $S=(K, \Gamma, F')$, then each $F'_{\sigma,\tau}$ is a unit of each K_P, so we have the representation (3) by the usual formula of §2, (7) for the $\mathfrak{p}$-component of a crossed product. Conversely, assume (3) for every $\mathfrak{p}$ involved in M. The algebra S has some crossed product representation $S=(K, \Gamma, \overline{F})$ with a factor set $\overline{F}$ not necessarily relatively prime to M. Then $S_\mathfrak{p}$ can be computed (by §2, (7)) as

$$S_\mathfrak{p} \sim (K_P/k_\mathfrak{p}, \Delta, \overline{F}_{\xi,\eta}) \sim (K_P/k_\mathfrak{p}, \Delta, E_{\xi,\eta})$$

where $E_{\xi,\eta}$ denotes the given factor set of P-adic units and $\Delta=\Delta(P)$. Since similar algebras arise only from similar factor sets, one has

$$\overline{F}_{\xi,\eta} \sim E_{\xi,\eta}, \quad \text{or} \quad \overline{F}_{\xi,\eta} = (A_\xi A_\eta^\xi A_{\xi\eta}^{-1})E_{\xi,\eta},$$

from some vector A_ξ of elements in K_P. One may then compute the character function h associated with the factor set of principal ideals $(\overline{F})$. By Lemma 5.6, putting $m=m_\mathfrak{p}$,

$$\begin{aligned} h_P(\zeta, (\overline{F})) &= (n/m)\sum_\eta V_P(\overline{F}_{\xi,\eta}) \\ &= (n/m)\sum_\eta V_P(A_\xi A_\eta^\xi A_{\xi\eta}^{-1}) \\ &= (n/m)\left[mV_P(A_\xi) + \sum_\eta V_P(A_\eta) - \sum_{\eta'} V_P(A_{\eta'})\right] \equiv 0 \pmod{n}. \end{aligned}$$

Therefore the $\mathfrak{p}$-primary part of $(\overline{F})$ has character 0 and so is similar to 1,

[18] Note that in the theorem we consider only finite prime divisors. For infinite unramified prime divisors $\mathfrak{p}$ there are no proper local division algebras. If $\mathfrak{p}$ is ramified then there is the algebra of all real quaternions. The condition F' relatively prime to M admits in the latter case the real quaternions. One finds that $F'_{1,\delta}F'_{\delta,\delta}<0$ (δ in the decomposition group of the ramified prime divisor) implies that $(K, \Gamma, F')_\mathfrak{p}$ is not similar to $k_\mathfrak{p}$.

which is to say that there are (p-primary) ideals $\mathfrak{B}_\sigma$ for which the similar factor set $(\overline{F})\cdot(T\mathfrak{B})$ is relatively prime to p. This can be done simultaneously for all the prime divisors p involved in M. One may then select another vector of ideals $\mathfrak{C}_\sigma$ such that $\mathfrak{C}_\sigma$ is relatively prime to M and $\mathfrak{B}_\sigma\mathfrak{C}_\sigma$ is a principal ideal (D_σ). We obtain then by the transformation

$$(\overline{F}) \sim (\overline{F})(T\mathfrak{B})(T\mathfrak{C}) = (\overline{F})(T\mathfrak{B}\mathfrak{C}) = (\overline{F})(TD)$$

a new factor set which is relatively prime to M. Thus $F' = \overline{F}\cdot TD$ is a new factor set of numbers which is similar to F and relatively prime to M. This gives the desired representation of S as $S = (K, \Gamma, F')$.

7. **Crossed characters and principal genera.** In considering transformation sets TA we shall repeatedly deal with the group of those vectors A_σ for which the transformation set TA is 1. Following a terminology suggested by A. H. Clifford, we call these vectors "*crossed characters*" of Γ. Specifically, let H be any multiplicative abelian group with Γ as a group of operators. Then a crossed character of Γ in H is any function $U(\sigma)$, with values $U(\sigma)$ in H for each σ in Γ, which satisfies the identity

(*7.1) $U(\sigma\tau) = U(\sigma)[U(\tau)]^\sigma$ (U a crossed character).

The vector group of all such crossed characters we call UH. From a fixed element H of the given group one may trivially obtain a crossed character $U(\sigma) = H^{1-\sigma}$. We call this a *unit character*, while two characters are *associates* if their quotient is a unit character. With the notation

(*7.2) UH = the group of all crossed characters in H,

(*7.3) $H^{1-\sigma}$ = the group of unit characters = the group of vectors $\{H^{1-\sigma}$, all σ in $\Gamma\}$,

the index $[UH : H^{1-\sigma}]$ will measure the number of classes of associated crossed characters of Γ in H.

A "*principal genus theorem*" is an assertion that every crossed character, under certain conditions, is a unit character. The principal genus theorem for ideals [33] is

THEOREM 7.1. *Every crossed character of the Galois group Γ in the group $\mathfrak{A}$ of all ideals of K is a unit character. The same conclusion holds for crossed characters in the group $\mathfrak{A}'$ of all ideals relatively prime to M.*

Proof. If $\mathfrak{C}_\sigma$ is a vector of ideals, the theorem asserts that $T\mathfrak{C} = 1$ implies $\mathfrak{C}_\sigma = \mathfrak{B}^{1-\sigma}$ for some ideal $\mathfrak{B}$. Let $\mathfrak{B}$ denote the greatest common divisor $\sum\mathfrak{C}_\sigma$ of the given ideals. Then one computes that $\mathfrak{B}^\sigma\mathfrak{C}_\sigma = \mathfrak{B}$ ([13, p. 127]). One may write $\mathfrak{B} = \mathfrak{B}'\mathfrak{P}$, where $\mathfrak{B}'$ is relatively prime to M and $\mathfrak{P}$ contains only prime factors of M. If each $\mathfrak{C}_\sigma$ is relatively prime to M, then

$$\mathfrak{P}^{1-\sigma} = \mathfrak{B}^{1-\sigma}/\mathfrak{B}'^{1-\sigma} = \mathfrak{C}_\sigma/\mathfrak{B}'^{1-\sigma}$$

is relatively prime to M for every σ, so that $\mathfrak{P}^{1-\sigma}=1$. Then $\mathfrak{C}_\sigma=\mathfrak{B}^{1-\sigma}=\mathfrak{B}'^{1-\sigma}$, so $\mathfrak{B}'$ is the desired ideal in the group $\mathfrak{A}'$.

There is also a similar theorem for numbers (Noether's principal genus theorem([19])).

THEOREM 7.2. *Every crossed character of* Γ *in the group of all nonzero numbers of* K *is a unit character*([20]).

LEMMA 1. *Given a factor set* $F=TB$, *with* F *but not necessarily* B_σ *relatively prime to* M, *there exists a vector* B'_σ *relatively prime to* M *for which* $F=TB'$.

Proof. Write the principal ideals (B_σ) as $(B_\sigma)=\mathfrak{A}_\sigma\mathfrak{C}_\sigma$, $\mathfrak{A}_\sigma$ relatively prime to M, where $\mathfrak{C}_\sigma$ involves only those prime ideals p of K which are involved in M. Then $(F)=T(B)=T\mathfrak{A}\cdot T\mathfrak{C}$. Since (F) does not involve any p of M, the transformation set $T\mathfrak{C}$ has no prime factors in common with either (F) or $T\mathfrak{A}$. Hence $T\mathfrak{C}=1$, and $\mathfrak{C}_\sigma=\mathfrak{B}^{1-\sigma}$, by the principal genus theorem. Choose $\mathfrak{A}'$ such that $\mathfrak{B}\mathfrak{A}'$ is a principal ideal (D). Then

$$(B_\sigma/D^{1-\sigma}) = (B_\sigma)/(D)^{1-\sigma} = \mathfrak{A}_\sigma\mathfrak{C}_\sigma/\mathfrak{B}^{1-\sigma}\mathfrak{A}'^{1-\sigma} = \mathfrak{A}_\sigma\mathfrak{A}'^{1-\sigma},$$

so that $B'_\sigma=B_\sigma/D^{1-\sigma}$ is, like $\mathfrak{A}_\sigma$, relatively prime to M. Furthermore,

$$TB' = TB/TD^{1-\sigma} = TB = F,$$

so the vector B'_σ has the desired properties.

In order to clarify the meaning of the principal class (F'') of ideal factor sets which appears in our basic index §6 (1), we now quote Noether's generalized principal genus theorem [33, 13]:

THEOREM 7.3. *If the vector* $\mathfrak{C}_\sigma$ *of ideals (not necessarily relatively prime to* M*) yields a transformation set* $T\mathfrak{C}$ *which lies in the principal class* (F'') *of ideal factor sets, then there exists an ideal* $\mathfrak{B}$ *of* K, *such that* $\mathfrak{B}^{1-\sigma}$ *and* $\mathfrak{C}_\sigma$ *lie in the same ideal class, for every* σ *in* Γ. *In other words, there exists a vector of principal ideals* (B_σ) *such that*

$$\mathfrak{C}_\sigma = (B_\sigma)\mathfrak{B}^{1-\sigma} \qquad (\textit{all } \sigma \textit{ in } \Gamma). \tag{1}$$

If $\mathfrak{C}_\sigma$ *is relatively prime to* M, $\mathfrak{B}$ *and* B_σ *may be chosen relatively prime to* M. *In other words, we have the implication*

$$T\mathfrak{C} = (F'') \rightarrow \mathfrak{C}_\sigma = (B_\sigma)\mathfrak{B}^{1-\sigma}. \tag{2}$$

Proof([21]). Let $T\mathfrak{C}=(F'')$. By definition (*6.2) of F'', the algebra S

([19]) Principal genus theorem "im Minimalen" [33].

([20]) The corresponding assertion for the group A' of all numbers relatively prime to M is false. If $p=(P_1\cdots P_g)^e$ is ramified in K/k, then there is an element B with $(B)=P_1\cdots P_g\mathfrak{C}$, $\mathfrak{C}$ relatively prime to M. The vector $C_\sigma=B^{1-\sigma}$ is relatively prime to M, but $C_\sigma=B'^{1-\sigma}$ can be shown to be impossible.

([21]) We repeat a proof here to check the provisions "relatively prime to M" which are not present to Noether's original theorem.

$=(K, \Gamma, F'')$ has $S_p \sim 1$ for any p involved in M. On the other hand, for a prime divisor p not involved in M, we may compute the invariant

$$\mu_p(S) = \mu_p(F'') = \mu_p(T\mathfrak{C}) \equiv 0 \pmod{n},$$

according to Theorem 5.4 and Lemma 5.1. Hence $S_p \sim 1$ for all p. This means that $S \sim 1$, so the factor set for S is $F'' = TB$, where the elements B_σ may be chosen relatively prime to M, as in Lemma 7.1. Thus $T\mathfrak{C} = (F'') = (TB)$, $T[\mathfrak{C}_\sigma(B_\sigma^{-1})] = 1$, so $\mathfrak{C}_\sigma(B_\sigma^{-1}) = \mathfrak{B}^{1-\sigma}$, by the principal genus theorem for ideals. This gives the result (2).

8. **Cyclic analogues to factor sets.** It has long been recognized that properties of factor sets provide parallels to the properties of the numbers which appear in the usual class field theory for cyclic fields Z/k. The parallel is as follows (we denote by λ a generator of the cyclic group of Z/k):

CYCLIC Z/k	ARBITRARY NORMAL K/k
1. A normal simple algebra S/k with maximal subfield Z (or K) may be represented as	
a cyclic algebra	a crossed product
$S = (Z, \lambda, a)$,	$S = (K, \Gamma, F)$.
2. This algebra is then determined by	
a, a number,	F, a factor set.
3. The associativity of the products in the algebra gives the condition	
a in k,	F satisfies the associativity conditions of §1, (4).
4. The algebra S is a total matric algebra if and only if	
a is a relative norm in Z/k:	F is a transformation set:
$a = N_{Z/k}C$,	$F = TA$.
5. The algebra S is a total matric algebra if and only if $S_p \sim k_p$ for all p; that is, if and only if, for each p,	
$a = N_p C_p$	$F \cap \Delta = T_\Delta A_P$
where N_p is the relative norm in Z_P/k_p,	where T_Δ is a transformation set for $\Delta = \Delta(P)$, A_P in K_P.
6. If B is a number of Z (B_σ a vector of K) then	
$N_{Z/k}B = 1$,	$TB_\sigma = 1$

holds if and only if, for some C in Z (or in K),

$$B = C^{1-\lambda} \qquad B_\sigma = C^{1-\sigma} \quad \text{(all } \sigma \text{ in } \Gamma\text{)}$$

(Hilbert's norm theorem), (minimal principal genus theorem).

7. A conductor

$$c(Z/k), \qquad C(K/k)$$

(see §26 below).

8. The principal genus consists of all ideals $\mathfrak{B}$ (ideal vectors $\mathfrak{B}_\nu$) relatively prime to the conductor c (to M) such that

$$N_{Z/k}\mathfrak{B} = (\nu), \qquad T\mathfrak{B} = (F''),$$

where the essential condition[22] is that, for every p in $C(Z/k)$ or M,

$$(Z, \lambda, \nu)_p \sim k_p, \qquad (K, \Gamma, F'')_p \sim k_p.$$

In the cyclic case the precise results of the class field theory depend on the computation that

$$[\mathfrak{a}' : (\nu)N\mathfrak{A}'] = n$$

where $\mathfrak{a}'$ denotes the group of all ideals in k which are relatively prime to the conductor $c(Z/k)$. Since norms correspond to transformation sets, elements of k to factor sets, ideals in k to ideal factor sets, and norm residues to factor sets F'', one sees that the corresponding index in the general case will be the index

$$J = [F\mathfrak{A}' : T\mathfrak{A}'(F'')]$$

which has already appeared in Theorem 6.1 as our main index J.

Next we shall show briefly that the general index J specializes to $[\mathfrak{a}' : (\nu)N\mathfrak{A}']$ in the cyclic case. So suppose that the cyclic Galois group is generated by λ. We can normalize the factor set in the usual manner[23]

$$u_\lambda^n = u_\lambda u_\lambda \cdots u_\lambda = \prod_{i=0}^{n-1} F_{\lambda^i,\lambda}.$$

We apply this process to the various groups involved in J, so

$$(1) \qquad F\mathfrak{A}' \to \prod_{i=0}^{n-1} \mathfrak{F}_{\lambda^i,\lambda} = \mathfrak{B},$$

$$(2) \qquad (F'') \to \prod_{i=0}^{n-1} (F''_{\lambda^i,\lambda}) = (B),$$

---

(22) The condition on ν is equivalent to requiring that ν be a norm residue for $c(Z/k)$.

(23) See, for example [13, pp. 64–65]; as well as §2, (9) above.

$$(3) \qquad T\mathfrak{A}' \to \prod_{i=0}^{n-1} (\mathfrak{A}'_\lambda \mathfrak{A}'^{\lambda^i}_\lambda \mathfrak{A}'^{-1}_{\lambda^{i+1}}) = N_{Z/k}\mathfrak{A}'_\lambda$$

gives a homomorphism of the various groups involved in J upon those of the classical class field theory. We first remark that $\mathfrak{B}$ is an ideal $\mathfrak{b}$ of k, for $\mathfrak{B}$ is relatively prime to M and invariant under λ. Inversion of the normalization of factor sets proves that every ideal $\mathfrak{a}'$ in k can be obtained as a $\mathfrak{B}$ from some $F\mathfrak{A}'$, and that every $N\mathfrak{A}'$ has the form $N\mathfrak{A}'_\lambda$ for some $T\mathfrak{A}'_\lambda$ of (3). It remains to prove only that the norm residues ν are the elements B obtained from the principal class (F'') in (2). Assume that the prime divisors p involved in the module M are precisely those which appear in the conductor $c(Z/k)$ (i.e., precisely those ramified in Z/k). Then B of (2) is in k, is relatively prime to $c(Z/k)$, and is the normalized constant of the algebra $S=(Z,\lambda,B)=(K,\Gamma,F'')$. By definition of F'', $S_\mathfrak{p}=(Z,\lambda,B)_\mathfrak{p}\sim k_\mathfrak{p}$ for all p in $c(Z/k)$. Therefore B is a local norm for each such p, so is a norm residue for $c(Z/k)$. Conversely, any norm residue ν is relatively prime to M and determines an algebra (Z,λ,ν) which is relatively prime to M and so can be written, by the inverse of the normalization process, as (Z,Γ,F''). This completes the proof of the assertion that J specializes to the classical index in the cyclic case(24).

We propose to investigate how far the methods used for the cyclic case will carry in the calculation of $J=[F\mathfrak{A}':T\mathfrak{A}'(F'')]$.

CHAPTER II. INDICES FOR GROUPS OF ALGEBRAS

9. Group-theoretic principles. The usual computations for the group indices in the cyclic case involve a number of principles for transforming given group indices. These we now state for reference. It is customary to carry out these group reductions formally, without any indication of purpose. Rather than join in this type of obscurantism, we attempt to formulate some directions for such calculations. If R, S, T(25) are given abelian groups, the usual problem is to compute the index $[R:S]$ of some subgroup S in R.

The objective can be attained if R and S are groups determined by simple and explicit generators, so that the quotient group R/S can be described completely and its order $[R:S]$ computed. For example, the group $\mathfrak{A}$ of all ideals has a simple generation by prime ideals. In order that this *direct computation* be possible, it is necessary to change a given index to indices involving other, simpler groups.

The simplest case is the *introduction* of suitable *intermediate* groups. In the class field theory it is especially useful to introduce a subgroup from the

(24) For general abelian extensions K/k the general index obviously does not specialize to $[\mathfrak{A}':(\nu)N\mathfrak{A}']$. The fact that the latter factor group is isomorphic with Γ cannot be explained by referring to general factor sets. It is a consequence of the special structure of K over k.

(25) The letters R, S, T are used in this section without reference to their previous and later meaning.

base field corresponding to a given group in the extension K. (Thus the group of numbers $\neq 0$ of k is a subgroup of the group A for K.) In general, if $[R:T]$ is finite,

$$R \supset S \supset T \rightarrow [R:T] = [R:S] \cdot [S:T]. \tag{1}$$

A *characterization* of a given group in different terms is often an essential preparatory step to a reduction. Thus the principal genus theorem for ideals characterizes the vectors $\mathfrak{B}^{1-\sigma}$ as those vectors $\mathfrak{C}_\nu$ for which $T(\mathfrak{C}_\nu) = 1$.

A *reduction principle* may be applied if a given group is a composite of two groups R and S. Since the groups are abelian, this composite is

(*9.1) $R \cdot S$ = the set of all products rs, for r in R, s in S.

One has [23, p. 129],

$$[RS:S] = [R:R \cap S], \tag{2}$$

where $R \cap S$ denotes the intersection. This equality may be proved on the assumption that either one of the two given indices is finite. If the composite occurs as a subgroup $ST \subset R$, one may introduce a smaller subgroup S and write, using (1)

$$[R:ST] = [R:S] \cdot [ST:S]^{-1} = [R:S] \cdot [T:S \cap T]^{-1} \tag{3}$$

on the assumption that $[R:S]$ is finite, or that $[R:ST]$ and $[T:S]$ are both finite.

The *isomorphism principle* is the well-known description of the effect of a homomorphism upon suitable subgroups. If ϕ is a homomorphic map of R on R', while S' is a subgroup of R', the set $\phi^{-1}(S')$ composed of all elements of R mapped by ϕ into S' is a subgroup of R, and the quotient groups $R/\phi^{-1}(S')$ and R'/S' are isomorphic. Hence

$$[R:\phi^{-1}(S')] = [R':S'], \tag{4}$$

provided either index is known to be finite.

The *homomorphism principle* describes the similar computation for any index $[R:S]$ under a homomorphism ϕ mapping R into part of a group T. Let $\phi^{-1}(1)$ denote the set of all elements mapped by ϕ onto the identity 1 in T. Then

$$\phi^{-1}(1) \cap S = \text{all elements of } S \text{ mapped onto 1 by } \phi. \tag{5}$$

One has

$$[R:S] = [\phi(R):\phi(S)] \cdot [\phi^{-1}(1) \cap R:\phi^{-1}(1) \cap S], \tag{6}$$

provided $[R:S]$ is finite.

This principle will be applied to two homomorphisms occurring naturally in our theory. The first is the map

(7) $$\mathfrak{A} \to \mathfrak{A}^{1-\sigma} = \mathfrak{A}/\mathfrak{A}^{\sigma}$$

carrying an ideal $\mathfrak{A}$ into the vector with components $\mathfrak{A}^{1-\sigma}$. The second is the map

(8) $$\mathfrak{A}_{\sigma} \to T\mathfrak{A}_{\sigma} = T\mathfrak{A} = \mathfrak{A}_{\sigma}\mathfrak{A}_{\tau}^{\sigma}/\mathfrak{A}_{\sigma\tau}$$

under which a vector $\mathfrak{A}_{\sigma}$ determines its correspondent transformation set. The analogous homomorphisms in the cyclic theory are

$$\mathfrak{A} \to \mathfrak{A}^{1-\lambda}, \qquad \mathfrak{A} \to N_{Z/k}\mathfrak{A}.$$

These homomorphisms have the convenient property that they may be applied in either order, one after the other, with the result always the identity

$$N(\mathfrak{A}^{1-\lambda}) = (N\mathfrak{A})^{1-\lambda} = 1.$$

This situation is exploited in the famous Herbrand group reduction principle [23, pp. 130–131]. In the general case this cannot be done, for it is meaningless to apply (8) "followed by" (7). This is the root for some essential difficulties in our computations (see §§18–23).

Another important homomorphism is

(9) $$A \to (A)$$

mapping the group of nonzero numbers on the group of principal ideals. This homomorphism is often applied backwards, to reduce an index on principal ideals to one on numbers. Under this homomorphism, the group mapped onto the identity is exactly the group of units. This is the point at which the units are inserted into the computations.

LEMMA 1. *Let $R \supset S$ be multiplicative groups of numbers in K, while (R), (S) are the corresponding groups of principal ideals, generated by these numbers. If E is the group of units of K, then*

(10) $$[(R):(S)] = [RE:SE],$$

provided either of the indices concerned is finite. If $R \supset E$, one also has

(11) $$[(R):(S)] = [R:S][E:S \cap E]^{-1}$$

provided the index $[R:S]$ is finite.

The proof will illustrate the systematic application of the principles above. Apply the homomorphism $R \to (R)$ and observe that the subgroup carried into the identity is simply the group E. Hence by (6), one has

$$[R:S] = [(R):(S)][R \cap E:S \cap E],$$

which gives (11). On the other hand, the homomorphism $RE \to (RE) = (R)$ is

one in which all elements mapped on 1 lie in the subgroup E of SE. Thus the isomorphism principle of (4) applies to give (10).

For direct products, one has

LEMMA 2. *Let $R_1 \subset R$ and $S_1 \subset S$ be subgroups of the respective factors of a direct product $R \times S$. Then*

$$[R \times S : R_1 \times S_1] = [R : R_1] \cdot [S : S_1], \tag{12}$$

provided the indices on either side of this equation are known to be finite.

As an application, suppose that $R \supset S$ are groups of numbers of K, while R_σ (S_σ) is the group of vectors with components R_σ (S_σ) in R (S) for every σ in Γ. The group R_σ is thus the direct product of n groups R, where n is the order of Γ. Hence, if $[R:S]$ is finite,

$$[R_\sigma : S_\sigma] = [R:S]^n. \tag{13}$$

10. **Invariant ideal classes.** The basic index J can be transformed into a form involving the number of ideal classes of K invariant under the group Γ. The results are stated in Theorems 10.1 and 10.2, while the method is directed at successive applications of the basic homomorphisms $\mathfrak{A} \to \mathfrak{A}^{1-\sigma}$ and $\mathfrak{A}_\sigma \to T\mathfrak{A}$. We need the notation:

(*10.1) $\mathfrak{A}^{1-\sigma}$ = a vector $\{\mathfrak{A}^{1-\sigma}, \sigma$ all elements of $\Gamma\}$, for $\mathfrak{A}$ a fixed ideal,
(*10.2) E = the group of all units in K,
(*10.3) $H = [\mathfrak{A} : (A)]$ = the class number of K,
(*10.4) $h = [\mathfrak{a} : (a)]$ = the class number of k.

Consider the index $J = [F\mathfrak{A}' : T\mathfrak{A}'(F'')]$. Since the second group is composite, one applies the appropriate reduction of §9, (3), to get

$$J = [F\mathfrak{A}' : (F'')][T\mathfrak{A}' : T\mathfrak{A}' \cap (F'')]^{-1}.$$

This is valid because the second index is finite, as will appear in the course of the subsequent computations of this section. The second member of J arises from a homomorphic map $\mathfrak{A}_\sigma \to T\mathfrak{A}$, and the principal genus theorem states that the vectors carried into the subgroup $T\mathfrak{A}' \cap (F'')$ by this map are exactly the vectors of the form $\mathfrak{A}^{1-\sigma}(A_\sigma)$, with $\mathfrak{A}$ and A_σ relatively prime to M. Hence the isomorphism principle yields

$$J = [F\mathfrak{A}' : (F'')][\mathfrak{A}_\sigma' : \mathfrak{A}'^{1-\sigma}(A')_\sigma]^{-1}. \tag{1}$$

Here the vector groups occurring on the right are (temporary notation):

$\mathfrak{A}_\sigma'$ = the group of all vectors $\mathfrak{A}_\sigma$ with components relatively prime to M,
$\mathfrak{A}'^{1-\sigma}$ = the group of all vectors of the form $\mathfrak{A}'^{1-\sigma}$, for an $\mathfrak{A}'$ relatively prime to M,
$(A')_\sigma$ = the group of all vectors of principal ideals generated by numbers A' relatively prime to M.

The composite appearing in the second index again suggests a reduction by (3) in §9,

$$(2) \qquad [\mathfrak{A}_s' : \mathfrak{A}'^{1-\sigma}(A')_s] = [\mathfrak{A}_s' : (A')_s][\mathfrak{A}'^{1-\sigma} : \mathfrak{A}'^{1-\sigma} \cap (A')_s]^{-1}.$$

This step (and the previous one) is valid if the index $[\mathfrak{A}_s' : (A')_s]$ on the right is finite. This it is, for such an index of two vector groups reduces, as in Lemma 2 of §9, to[26]

$$[\mathfrak{A}_s' : (A')_s] = [\mathfrak{A}' : (A')]^n = H^n.$$

But the second member of (2) suggests the homomorphism $\mathfrak{A} \to \mathfrak{A}^{1-\sigma}$. For the elements carried by this homomorphism into the subgroup $\mathfrak{A}^{1-\sigma} \cap (A)_s$ we use the letter $\mathfrak{D}$:

(*10.5) $\mathfrak{D}$ = all ideals in K with $\mathfrak{D}^{1-\sigma}$ principal for every σ in Γ.

For the moment we also denote by $\mathfrak{D}'$ the ideals of $\mathfrak{D}$ relatively prime to M. These ideals $\mathfrak{D}$ might be called ideals in "invariant classes," for by definition $\mathfrak{D}^\sigma$ and $\mathfrak{D}$ lie in the same ideal class, for every σ. Under this homomorphism, (2) becomes

$$(3) \qquad [\mathfrak{A}_s' : \mathfrak{A}'^{1-\sigma}(A')_s] = H^n[\mathfrak{A}' : \mathfrak{D}']^{-1}.$$

The quotient $H[\mathfrak{A}' : \mathfrak{D}']^{-1}$ involved here suggests the elimination of the group $\mathfrak{A}'$,

$$(4) \qquad H[\mathfrak{A}' : \mathfrak{D}']^{-1} = [\mathfrak{A}' : (A')][\mathfrak{A}' : \mathfrak{D}']^{-1} = [\mathfrak{D}' : (A')].$$

In the latter index, one may drop the condition "relatively prime to M." Specifically, any ideal $\mathfrak{D}$ generates an ideal class which contains some ideal $\mathfrak{B}'$ relatively prime to M, so $\mathfrak{D} = \mathfrak{B}'(A)$. Here $\mathfrak{B}'^{1-\sigma}$, like $\mathfrak{D}^{1-\sigma}$, must be principal, so that $\mathfrak{B}'$ is in $\mathfrak{D}'$, and $\mathfrak{D}$ is the group join $\mathfrak{D}'(A)$. By the reduction principle

$$[\mathfrak{D} : (A)] = [\mathfrak{D}'(A) : (A)] = [\mathfrak{D}' : (A) \cap \mathfrak{D}'].$$

The intersection $(A) \cap \mathfrak{D}'$ is just the group of principal ideals (A'), so

$$(5) \qquad [\mathfrak{D} : (A)] = [\mathfrak{D}' : (A')].$$

A combination of the results of (1), (3), (4), and (5) yields

THEOREM 10.1. *If H is the class number of K, while $[\mathfrak{D} : (A)]$ is the number of invariant ideal classes of K (see (*10.5)), then the basic index is*

$$(6) \qquad J = [F\mathfrak{A}' : (F'')][\mathfrak{D} : (A)]^{-1}H^{1-n},$$

*where the first factor gives the number of ideal factor sets relatively prime to M modulo the "principal class" (F'') of (*6.2).*

[26] Note that the class number H may be computed with a restriction to ideals relatively prime to M (any ideal class contains an ideal relatively prime to any given module).

To study the index $[\mathfrak{D}:(A)]$ appearing above, we use the definition of $\mathfrak{D}$, which involves the mapping of $\mathfrak{D}$ on the vector $\mathfrak{D}^{1-\sigma}$ of principal ideals. The subgroup carried into the identity by this mapping is the *group of invariant ideals* $\mathfrak{J}$,

$$\mathfrak{J} = \text{all ideals of } K \text{ with } \mathfrak{J} = \mathfrak{J}^{\sigma}, \text{ for every } \sigma \text{ in } \Gamma.$$

The corresponding subgroup of invariant principal ideals is (Δ), where

(*10.6) $\Delta = \text{all numbers} \neq 0 \text{ of } K \text{ with } (\Delta^{1-\sigma}) = 1, \text{ for every } \sigma.$

According to the homomorphism principle we get,

$$[\mathfrak{D}:(A)] = [\mathfrak{D}^{1-\sigma}:(A^{1-\sigma})][\mathfrak{J}:(\Delta)]. \tag{7}$$

The group of principal ideals (Δ) certainly includes all principal ideals (a) of k. The introduction of this subgroup in the second factor of (7) yields

$$[\mathfrak{J}:(\Delta)] = [\mathfrak{J}:(a)][(\Delta):(a)]^{-1}, \tag{8}$$

provided the first index is finite. To see this, use the subgroup of ideals $\mathfrak{a}$ in k, for which

$$[\mathfrak{J}:(a)] = [\mathfrak{J}:\mathfrak{a}][\mathfrak{a}:(a)].$$

The second factor is the (finite) class number h, while the first measures the number of invariant ideals $\mathfrak{J}$ which are not extensions of ideals in k. But an invariant ideal $\mathfrak{J}$ involves with each prime factor P every conjugate P^{σ}. If P is unramified in K/k, the product $P_1P_2 \cdots P_g$ of all conjugates P_i of P is a prime ideal p of k, consequently is in the group $\mathfrak{a}$. Hence any $\mathfrak{J}$ is congruent modulo $\mathfrak{a}$ to an invariant ideal $\mathfrak{J}$ involving only prime factors P ramified in K/k. If p is a ramified ideal in K/k, with

$$p = (P_1 \cdots P_g)^e, \qquad e = e_p, \tag{9}$$

then the invariant ideals involving only factors of p must all be of the form $(P_1 \cdots P_g)^i$. For these ideals we get a complete set of representatives modulo $\mathfrak{a}$ if i ranges from 0 to $e-1$, for $(P_1 \cdots P_g)^e = p$ is an ideal in $\mathfrak{a}$. Combining the effects due to the different ramified primes (according to the direct product of the partial groups $\mathfrak{J}$) one finds

$$[\mathfrak{J}:\mathfrak{a}] = \prod e_p \qquad \text{(over all finite } p \text{ of } k).$$

Therefore the first factor on the right of (8) is indeed finite, and the result is

$$[\mathfrak{J}:(\Delta)] = h(\prod e_p)[(\Delta):(a)]^{-1}. \tag{10}$$

The index $[(\Delta):(a)]$ may be shifted to one involving only numbers. Since every unit E is by definition (*10.6) a number Δ, one has, by equation (10) of §9,

$$[(\Delta):(a)] = [\Delta:aE].$$

To exploit the definition of Δ one clearly must use the homomorphism $\Delta \to \Delta^{1-\sigma}$. In this homomorphism the only elements mapped on the identity are in a, so the isomorphism principle yields

$$[(\Delta):(a)] = [\Delta^{1-\sigma}:E^{1-\sigma}].$$

The vector $\Delta^{1-\sigma}$ can be characterized as a vector of units E_σ which can be obtained as $E_\sigma = A^{1-\sigma}$ for a number $A = \Delta$. By the principal genus theorem for numbers, the vector E_σ has this form if and only if $TE_\sigma = TE = 1$, i.e., if and only if E_σ is a crossed character $U(\sigma)$ (see §7).

(*10.7) UE = the group of all crossed characters of Γ in the group of units = all vectors of units $U(\sigma)$ with $U(\sigma)[U(\tau)]^\sigma = U(\sigma\tau)$.

Then

$$[(\Delta):(a)] = [UE:E^{1-\sigma}]. \tag{11}$$

THEOREM 10.2. *The number of invariant ideal classes of K is*

$$[\mathfrak{D}:(A)] = [\mathfrak{D}^{1-\sigma}:(A^{1-\sigma})][UE:E^{1-\sigma}]^{-1} h \prod e_p, \tag{12}$$

where $[UE:E^{1-\sigma}]$ denotes the number of non-associated crossed characters in E, h is the class number of the base field k, and the product $\prod e_p$, taken over all finite prime divisors p, uses

(*10.8) *e_p = the ramification order of the prime ideal p in K/k.*

The crossed characters of units are closely related by (11) to the principal ideal theorem [30]. The group (Δ) of principal invariant ideals includes the group $\mathfrak{a}_K$ of those ideals in k which become principal ideals in K, so that

$$[\mathfrak{a}_k:(a)] \leqq [UE:E^{1-\sigma}]. \tag{13}$$

In case K/k is unramified, the above computation of $[\mathfrak{J}:\mathfrak{a}]$ shows that all (Δ) lie in $\mathfrak{a}_K$, so that this inequality becomes an equation. We state the result as

THEOREM 10.3. *The number $[\mathfrak{a}_K:(a)]$ of ideal classes of k which become principal in K is at most equal to the number of classes of associated crossed characters of Γ in the group E of units. If K/k is unramified, these two numbers are equal.*

The equation (11) in reality involves an isomorphism $(\Delta)/(a) \simeq UE/E^{1-\sigma}$. As Moriya ([30]) has done in the cyclic case, one may call those crossed characters, which in this isomorphism correspond to elements in the subgroup $\mathfrak{a}_K/(a)$, crossed characters of the "first kind." This allows an obvious restatement of Theorem 10.3.

11. **Reduction to unit factor sets.** A further reduction leads now to the formula for J given in Theorem 11.2 below. We use the notation

(*11.1) FE = the group of factor sets of units,

(*11.2) TE = the group of all transformation sets of units.

The index $[\mathfrak{D}^{1-\sigma}:(A^{1-\sigma})]$ of Theorem 10.2 may be shifted from ideals to numbers if one observes that by definition (*10.5) each $\mathfrak{D}^{1-\sigma}$ is a principal ideal (θ_σ). Let θ_σ temporarily denote any vector of numbers so obtainable,

$$(\theta_\sigma) = \mathfrak{D}^{1-\sigma}, \qquad \text{for some } \mathfrak{D} \text{ and all } \sigma.$$

Surely a vector of units E_σ is such a vector, for each (E_σ) is 1. Therefore the principal ideal shift of §9, applied to these vector groups, yields

$$[\mathfrak{D}^{1-\sigma}:(A^{1-\sigma})] = [(\theta_\sigma):(A^{1-\sigma})] = [\theta_\sigma:A^{1-\sigma}E_\sigma].$$

Now apply the homomorphism $\theta_\sigma \to T\theta_\sigma = T\theta$. The elements mapped on 1 by this homomorphism are all in the group of vectors $A^{1-\sigma}$, according to the minimal principal genus theorem (see §7). Hence the isomorphism principle yields

$$[\mathfrak{D}^{1-\sigma}:(A^{1-\sigma})] = [\theta_\sigma:A^{1-\sigma}E_\sigma] = [T\theta:TE]. \tag{1}$$

We next investigate the group $T\theta$. By definition, $(\theta_\sigma)=\mathfrak{D}^{1-\sigma}$, so $T(\theta)=T\mathfrak{D}^{1-\sigma}=1$. Each element of the factor set $T\theta$ is thus a unit of K, so $T\theta$ is contained in the group FE of factor sets of units. On the other hand, $T\theta$ is a transformation set of numbers, hence the corresponding crossed product algebra $S=(K, \Gamma, T\theta)$ is a total matric algebra. Since each p-adic component $S_\mathfrak{p}$ of this algebra is then similar to $k_\mathfrak{p}$, it follows that $T\theta$ is one of the factor sets F'' considered in our basic index.

We now assert that these two groups FE and F'' not only both contain $T\theta$, but that their intersection is $T\theta$,

$$FE \cap F'' = T\theta. \tag{2}$$

For, let F be a factor set in the intersection(27), and consider the invariants of the crossed product $(K, \Gamma, F)=S$. If $\mathfrak{p}$ is ramified in K/k, the assumption F in F'' means that $S_\mathfrak{p} \sim 1$, hence that the corresponding invariant $\mu_\mathfrak{p} \equiv 0 \pmod{n}$. If $\mathfrak{p}$ is finite and not ramified in K/k, the invariant $\mu_\mathfrak{p}(S)$ may be computed by the explicit formula of §2, (10). Since all the $F_{\sigma,\tau}$ of the factor sets are units in K, the invariant turns out to be $\equiv 0 \pmod{n}$. If $\mathfrak{p}$ is infinite and unramified the invariant is also $\equiv 0 \pmod{n}$. As the algebra S is completely determined (up to similarity) by its invariants, this proves $S \sim 1$. Therefore the factor set F is a transformation set $F=TA_\sigma=TA$. The principal ideals $T(A_\sigma)=(TA_\sigma)=(F_{\sigma,\tau})$ are then all equal to 1. The principal genus theorem for ideals then asserts that $(A_\sigma)=\mathfrak{D}^{1-\sigma}$ for some ideal $\mathfrak{D}$. Therefore the vector A_σ is one of the vectors θ_σ, $F=TA$ lies in the group $T\theta_\sigma$, and (2) is established.

Introducing this expression in the index (1), we have

THEOREM 11.1. *The first index of Theorem* 10.2 *is*

(27) This proof is essentially due to E. Noether [33].

$$(3) \qquad [\mathfrak{D}^{1-\sigma}:(A^{1-\sigma})] = [FE \cap F'':TE].$$

The same intersection group may be extracted from the first factor of J, as found in Theorem 10.1. In $[F\mathfrak{A}':(F'')]$ insert the intermediate group (F') of ideal factor sets (see (*3.2)):

$$[F\mathfrak{A}':(F'')] = [F\mathfrak{A}':(F')][(F'):(F'')].$$

Here the second index may be shifted to numbers, after the manner of §9, (11), with the results

$$(4) \qquad [F\mathfrak{A}':(F'')] = [F\mathfrak{A}':(F')][F':F''][FE:F'' \cap FE]^{-1}$$

which is valid provided $[F':F'']$ is finite.

When this result is inserted in the expression (6) for J in Theorem 10.1 and combined with the results of Theorems 10.2 and 11.1, we find a formula for J, the denominator of which involves parts of (4), as

$$[FE:F'' \cap FE][FE \cap F'':TE] = [FE:TE].$$

All told, one has

THEOREM 11.2. *If the factor set index $[F':F'']$ is finite*[25] *(see (*6.1)), then*

$$(5) \qquad \begin{aligned} J &= [F\mathfrak{A}':(F')]H^{1-s}h^{-1}[F':F''](\prod e_\mathfrak{p})^{-1}J(E), \\ J(E) &= [UE:E^{1-\sigma}][FE:TE]^{-1}, \end{aligned}$$

where the product is taken over the ramification orders $e_\mathfrak{p}$ of all finite primes, where H and h are class numbers, and $[F\mathfrak{A}':(F')]$ is the number of ideal factor sets relatively prime to M modulo the principal ideal factor sets.

The group E of units in K appears in (5), as the quotient $J(E)$: the number of classes of crossed characters of E divided by the number of group extensions of E. In the sequel we turn to the separate investigation of the terms in (5). We first compute $[F':F'']$, proving it finite, then reduce the index $[F\mathfrak{A}':(F')]$ in terms of a certain group-theoretic "deficiency" index, and finally devote a chapter to the index $J(E)$.

12. **Local algebras with factor sets of units.** The study of factor sets F'' for algebras unramified at the divisors of M will subsequently be reduced to questions on local algebras $S_\mathfrak{p}$ with factor sets of $\mathfrak{p}$-adic units. For any local algebra $S_\mathfrak{p}$ split by $K_P/k_\mathfrak{p}$ we consider the possible crossed product representation

$$(1) \qquad S_\mathfrak{p} = (K_P/k_\mathfrak{p}, \Delta(P), F)$$

in which $F = F_{\zeta,\eta}$ denotes now a factor set defined for the elements ζ, η in the Galois group $\Delta(P) = \Delta$ of $K_P/k_\mathfrak{p}$. Since the invariant μ of this algebra is

[25] The finiteness $[F':F'']$ will be proved later, in Theorem 13.1.

an element of an additive cyclic group, the class group of these algebras is cyclic of order $m_p = m = [K_P : k_p]$. Consider now only those factor sets F which consist of P-adic units.

THEOREM 12.1. *The p-adic algebra classes S_p which have a crossed product representation* (1) *with a factor set FE_P consisting of P-adic units $E_{\xi,\eta}$ form a cyclic group of order e_p, where e_p is the ramification order of K_P/k_p.*

Proof. If K_p is unramified over k_p, $e = e_p = 1$, and the theorem follows readily. For $\Delta = \{\delta\}$ is then a cyclic group, so the invariant $\mu(S_p)$ can be computed as in (10) of §2. The result is a sum of terms $V_P[E_{\delta^i,\delta}]$ which are all zero because the P-adic order of a unit $E_{\delta^i,\delta}$ is zero. Since the invariant is zero, the algebra $S_p \sim k_p$, as asserted. In the general case the maximal unramified subfield W of K_P has over k_p a degree $f = m/e$. Any algebra S_p has as index a divisor of m, and the index is the same as the exponent of S_p in the group of algebra classes. Hence the index of the power S_p^e is a divisor of m/e. This means that S_p^e has W/k_p as splitting field, so that S_p^e is similar to some crossed product of W. By a formula due to Witt [41], one can explicitly calculate this representation of S_p^e. If Ω is the subgroup of Δ corresponding to W according to the Galois theory, the extension W/k_p has the factor group Δ/Ω as Galois group. If for each coset $\eta\Omega$ of this factor group a representative η' in Δ is selected, Witt showed that

$$S_p^e = (K_P, \Delta, FE_P)^e \sim (W/k_p, \Delta/\Omega, B) \tag{2}$$

where the factor set B consists of quantities given in terms of $FE_P = E_{\xi,\eta}$ as

$$B_{\xi\Omega,\eta\Omega} = \prod_{\omega} E^{\omega}_{\xi',\eta'} E_{\omega,\xi'\eta'} E^{-1}_{\omega,(\xi\eta)'} \qquad \text{(over all } \omega \text{ in } \Omega\text{)}.$$

For our purpose we need only note that if the original factor set FE_P consists of units, then this derived factor set will also consist of units. By the computation already made for the unramified case, the algebra $(W, \Delta/\Omega, B)$ of (2) is then similar to k_p, so that $S_p^e \sim k_p$. In the group of all algebras with factor sets of units any algebra thus has order at most e. This group must be cyclic by the remark at the beginning of this section. Consequently to complete the proof of the present theorem we need only show that some algebra with factor set of units has order at least e. In the whole (cyclic) group of all local algebras S_p there is one algebra of order $m = ef$. Hence it will suffice to show that the fth power of this algebra has a factor set of units. This is indeed the case, as we shall prove in

THEOREM 12.2. *Any factor set F for K_P/k_p has its fth power F^f similar to a factor set of units.*

Proof. The integer $f = m/e$ can be described as the inertial degree of K_p

over k_p. If the factor set F does not already consist of P-adic units, denote by $e_{\tau,\eta}$ the P-adic orders $V_P[F_{\tau,\eta}]$ of its elements. The associativity relations

$$F_{\tau,\eta}F_{\tau\eta,\xi} = F_{\eta,\xi}^{\tau}F_{\tau,\eta\xi},$$

if multiplied over all values of ζ (in Δ), yield for the e's the relations

$$b_\eta + b_\xi = me_{\eta,\xi} + b_{\eta\xi}. \tag{3}$$

Here b_η denotes the P-adic order $V_P(B_\eta)$ of a vector B_η, with

$$B_\eta = \prod_\tau F_{\tau,\eta}, \qquad C_\tau = \prod_\eta F_{\tau,\eta}.$$

For these two vectors the associativity relations, multiplied over all values of η, yield

$$C_\tau B_\xi = B_\xi^\tau C_\tau,$$

which is to say that $B_\xi^\tau = B_\xi$. The invariant element B_ξ must therefore lie in the subfield k_p, which implies in turn that its P-adic order b_ξ is a multiple of the ramification order e. Hence $b_\xi = ea_\xi$ for a suitable integer a_ξ, and (3) can be rewritten as

$$fe_{\eta,\xi} = a_\eta + a_\xi - a_{\eta\xi}.$$

If A_τ is a vector of elements of the respective orders $-a_\tau$, this means that

$$(F_{\eta,\xi})^{\prime}TA_\tau$$

will be a factor set each of whose elements has P-adic order zero (i.e., is a P-adic unit). In other words F' is similar to a factor set of P-adic units, as asserted.

If p is an infinite prime divisor, Theorem 12.1 still is valid, if we adopt the usual convention as to ramifications at infinity. If $K_P = k_p$, the extension K_P/k_p is, of course, unramified, and $e_p = 1$. If K_P is the field of complex numbers, k_p is that of real numbers, then K_P is ramified over k_p with ramification order $e_p = 2$. Theorem 12.1 then holds because the group of algebras S_p over k_p is then generated by the algebra of real quaternions.

13. **Algebras with factor sets relatively prime to** M. One of the partial indices in the formulas (5) of §11 is $[F' : F'']$, for the group F' of all factor sets of numbers relatively prime to M and the subgroup F'' consisting of those factor sets F' for which $(K, \Gamma, F'')_p \sim k_p$ whenever p is involved in M. To reduce this index, we apply the natural homomorphism carrying F into a crossed product

$$F \to S = (K, \Gamma, F). \tag{1}$$

According to Theorem 6.2 we map F' on the group S_s, defined by

S_* = all normal simple algebras S/k split by K/k such that, whenever p is involved in M, S_p has a crossed product representation $(K_P/k_p, \Delta(P), FE_P)$, with a factor set FE_P of P-adic units.

The subgroup F'' is, by its very definition, the group of those factor sets carried by the homomorphism (1) into algebras S' relatively prime to M. Hence, by the isomorphism principle[29],

$$[F' : F''] = [S_* : S']. \tag{2}$$

The index on the right can be computed in terms of the behaviour of local algebras S_p at the ramified prime divisors. Let us use the following notation:

(*13.1) $p_1, p_2, \cdots, p_t$ all (finite and infinite) prime divisors of k which are ramified in K,

(*13.2) e_i = the ramification order of a divisor P_i of p_i,

(*13.3) $m_i = [K_{P_i} : k_{p_i}]$ = the local p_i-degree of K/k.

Then $e_i \mid m_i$ and $m_i \mid n$ (n/m_i is the number of distinct prime factors of p).

THEOREM 13.1. *The index $[F' : F'']$ may be computed in terms of invariants of K/k as*

$$[F' : F''] = n^{-1}(n, J(\Gamma)n/e_1, \cdots, J(\Gamma)n/e_t) \prod_p e_p, \tag{3}$$

where the product is taken over all (finite and infinite) ramification divisors of K/k, while $J(\Gamma)$ is the least common multiple of the orders of the elements of the Galois group Γ of K/k.

For a proof we appeal to the local invariants of the algebras S_*. If p_i is a ramified prime divisor, the component S_p of an algebra in S_* has by definition a factor set of P_i-adic units. In the group of all algebra classes split by K_P/k_p, S_P consequently has an order which divides e_i (Theorem 12.1). The integral invariant of S at p_i thus equals[30] $\mu_i(S) = x_i(n/e_i)$, where x_i is an integer which is uniquely determined mod e_i. We map the group S_* on a vector group of these invariants,

$$X = (x_1, x_2, \cdots, x_t), \qquad nx_i = e_i\mu_i(S). \tag{4}$$

---

(29) This equation may also be viewed as follows. The homomorphism which carries a factor set F into an algebra S is in effect the reduction of the group of factor sets modulo the transformation sets. The group of algebras S_* obtainable from F' is therefore isomorphic with $F'/(TA \cap F')$. Similarly, S' is isomorphic with $F''/(TA \cap F'')$. But the intersections $TA \cap F'$ and $TA \cap F''$ are identical, for a transformation set TA which is relatively prime to M will necessarily determine a total matric algebra. Hence it will lie in F''. Consequently $[S_* : S'] = [F'/(TA \cap F') : F''/(TA \cap F')] = [F' : F'']$.

(30) For simplicity we write the local algebra S and the invariant μ with i instead of p_i as subscript.

For all the unramified prime divisors p which are involved in M we have $S_p \sim k_p$. Hence an algebra S of S_s belongs to the subgroup S' if and only if each $S_i \sim k_i$; that is, if and only if the vector X of invariants corresponding to S has the form

$$Y = (y_1 e_1, y_2 e_2, \cdots, y_t e_t),$$

with invariants $\mu_i(S) = y_i e_i (n/e_i) \equiv 0 \pmod{n}$. Therefore our index becomes

$$[F' : F''] = [S_s : S'] = [X : Y].$$

If we introduce the group of all vectors

$$Z = (z_1, z_2, \cdots, z_t) \quad \text{(each } z_i \text{ a rational integer)},$$

this index will become

$$[F' : F''] = [Z : Y]/[Z : X] = \prod_p e_p/[Z : X], \tag{5}$$

where the product is taken over all prime divisors of k.

It remains only to compute the index $[Z : X]$, which measures how many of the *a priori* conceivable sets of invariants z_i are possible for an actual algebra S. The only condition on the invariants of an algebra S is the sum relation (see [22]),

$$0 \equiv \sum_p \mu_p(S) \equiv \sum_{i=1}^{t} \mu_i(S) + \sum_q \mu_q(S) \equiv \sum_{i=1}^{t} x_i(n/e_i) + \sum_q \mu_q(S) \pmod{n}, \tag{6}$$

where q runs over all prime divisors distinct from $p_1, \cdots, p_t$. The invariants $\mu_q(S) = \mu_q$ can be considered as the invariants of an ideal algebra $\mathfrak{S}$ which is unramified at every p_i $(i = 1, 2, \cdots, t)$, but which otherwise has the same components as does S. We have

LEMMA 1. *An integer r can be the sum $\sum_p \mu_p(\mathfrak{S}')$ of the invariants of an ideal algebra $\mathfrak{S}'$ if and only if $J(\Gamma) r \equiv 0 \pmod{n}$.*

Proof. In establishing Theorem 4.1 we showed that any r satisfying the above condition is the invariant sum of a suitable ideal algebra $\mathfrak{S}'$. Conversely, an invariant $\mu_p(\mathfrak{S}')$ for an unramified p has by definition (see §2, (3)) the form sn/m_p, where s is an integer and m_p is the order of a corresponding Frobenius automorphism. Since $m_p \mid J(\Gamma)$ we have $J(\Gamma)\mu_p(\mathfrak{S}') \equiv 0$, and $J(\Gamma)\sum_p \mu_p(\mathfrak{S}') \equiv 0 \pmod{n}$.

LEMMA 2. *A set of integers x_i belongs to the group X of* (4) *if and only if $\sum_{i=1}^{t} x_i (J(\Gamma) n/e_i) \equiv 0 \pmod{n}$.*

Proof. The relation (6) characterizes the integers x_i by the condition that $-\sum_{i=1}^{t} x_i (n/e_i)$ be congruent, mod n, to the invariant sum of an ideal algebra $\mathfrak{S}'$. Lemma 1 now gives the result desired.

By the elementary theory of congruences it is now possible to find a basis for the group X. One need only select from those vectors in which the first $i-1$ components are zero a vector $(0, 0, \cdots, 0, x_{ii}, \cdots, x_{it})$ in X with a minimal ith component $x_{ii}>0$. The index $[Z:X]$ of (5) is then $\prod_{i=1}^{t} x_{ii}$. This product can be computed as n divided by the greatest common divisor

(*13.4) $n^* = (n, J(\Gamma)n/e_1, \cdots, J(\Gamma)n/e_t)$.

The formula (5) then becomes the assertion (3) of Theorem 13.1.

We pause to discuss the invariant n^* of (*13.4), which appears in the formula (3). Clearly $J(\Gamma) \mid n^* \mid n$. In the definition (*13.4) one may omit any e_i which divides $J(\Gamma)$. Suppose in particular that some p_i has no higher ramification([31]). Then the ramification order e_i is the order of the inertial group of $P_i \mid p_i$, which is cyclic, so certainly $e_i \mid J(\Gamma)$. The formula (*13.4) thus need include only the prime divisors p_i with higher ramification.

THEOREM 13.2. *The invariant of* (*13.4) *satisfies* $n^*=n$ *if and only if the ramification order* e *of every* p *in* K/k *is a divisor of* $J(\Gamma)$. *In particular,* $n^*=n$ *whenever* $J(\Gamma)=n$, *or whenever there is no prime ideal* p *in* k *with higher ramification in* K.

THEOREM 13.3. *Let* l^r *be the exact power of a rational prime* l *which divides the degree* $n=[K:k]$, *while* l^s *is the largest power of* l *which occurs as the order of an element in the Galois group* Γ. *Among the prime ideal divisors* P *of* l *in* K *select one whose Hilbert ramification group* V_1 *has as large an order* l^r *as possible. Then the exact power* l^u *of* l *dividing* n^* *is given by*

$$u = r \qquad (\text{if } r < s),$$

$$u = r - (r - s) \qquad (\text{if } r \geqq s).$$

Proof. This theorem follows at once from (*13.4), for the ramification order e of a prime ideal which divides the rational prime l has the form $e=e_0 l^r$, where $(e_0, l)=1$, while l^r is the order of the first Hilbert ramification group.

The inequality $n^*<n$ can actually arise, even though an earlier summary of these results ([29]) was based on the assumption $n^*=n$. For a simple explicit example, let k be the field R of all rational numbers, while $K=R(6^{1/2}, 7^{1/2})$ is a quartic field with the four group as Galois group. In each of the three quadratic subfields $R(6^{1/2})$, $R(7^{1/2})$, and $R(42^{1/2})$, the rational prime 2 is ramified, since each field has an even discriminant. It follows that the ideal (2) is totally ramified in K/R, for in any other event, K would have a quadratic subfield which is an inertial or decomposition field for (2), counter to the observation above. Since (2) has ramification order 4, while $J(\Gamma)=2$, we conclude by Theorem 13.2 that $n^*=2$, $n=4$.

A more general construction is embodied in

([31]) Each prime factor P_i of p_i has its first ramification group $V_1=1$. See [20, Part Ia, p. 70].

THEOREM 13.4. *Let l be a rational prime which is not totally decomposed in an algebraic number field k. Then k has an abelian extension K with a degree $n=l^l \geqq l^3$ such that the invariant n^* is at most l^2.*

Proof. The local class field theory [9] may be used to construct an abelian extension of type $(l, l, \cdots, l)$ with sufficiently complicated ramifications. By hypothesis, l has in k some prime factor p with a local degree $n_p=[k_p:R_l]\geqq 2$. Consider the corresponding group k_p^*/k_p^{*l}. This group can be explicitly expressed by using a basis of the units in k_p. In the regular case, when k_p contains no primitive lth roots of unity, the group k_p^*/k_p^{*l} is abelian of type $(l, l, \cdots, l)$, with n_p+1 generators. In the irregular case, when k_p contains all lth roots of unity, the group k_p^*/k_p^{*l} is abelian of type $(l, l, \cdots, l)$ with n_p+2 generators. (See [18] and [24].) In either event the existence theorem of local class field theory asserts that k_p has an abelian extension A_p with a Galois group isomorphic with k_p^*/k_p^{*l}. The degree of this extension is at least $l^{n_p+1}\geqq l^3$, since by hypothesis $n_p\geqq 2$. In A_p/k_p, the Galois group modulo the inertial group is cyclic, consequently it has order at most l. By Grunwald's existence theorem [16], the local extension A_p/k_p can be obtained from infinitely many abelian extensions A/k with the same Galois groups as A_p/k_p. The invariant n^* of (7) is then $n^*\leqq(n, J(\Gamma)n/e)\leqq(n, ll)=l^2$, as asserted.

These examples may be combined to give the following conclusion:

THEOREM 13.5. *Over any algebraic number field k there is a normal (abelian) extension K with*(32) $n^*<n$.

14. **Algebras with given local components.** The questions raised in §13 as to the existence of local algebras with specified local components suggest an analogous inquiry: when is a given local algebra a component of an actual algebra? This leads to a certain index $[\mathfrak{S}^{(p)}:S^{(p)}]$, analogous to the main index $[\mathfrak{S}':S']$, where

(*14.1) $\mathfrak{S}^{(p)}$ = the group of classes of normal simple algebras $S^{(p)}$ split by K_P/k_p,

(*14.2) $S^{(p)}$ = the group of those algebras $\mathfrak{S}^{(p)}$ which are components of some S split by K/k.

If p is unramified, the invariant of $\mathfrak{S}^{(p)}$ has the form $\mu_p(\mathfrak{S}^{(p)})\equiv x(n/m_p) \pmod n$. By the Tschebotareff density theorem we know that $m_p=m_q$ for at least one other unramified prime ideal q of k. There is thus a local algebra S_q with an invariant $\mu_q(S_q)\equiv -x(n/m_q) \pmod n$. These two invariants add up to $\mu_q(S_q)+\mu_p(\mathfrak{S}^{(p)})\equiv 0 \pmod n$, so there is an actual algebra S with $\mathfrak{S}^{(p)}$ and $\mathfrak{S}_q$ as its only non-trivial components. Therefore $S^{(p)}=\mathfrak{S}^{(p)}$ for any unramified p.

Consider next the t ramified prime divisors p_i with associated decomposi-

(32) For the existence of suitable primes l consult the density theorems. See [20, Part II].

tion groups of order([33]) m_i. The invariants of an actual algebra S must satisfy the relation

$$\mu_i(S) + \sum_{j\neq i,\, j=1}^{t} \mu_j(S) + \sum_{q} \mu_q(S) \equiv 0 \pmod{n} \qquad (q \neq p_1, \cdots, p_t).$$

The terms here can be arbitrary multiples of n/m_i, n/m_j, and n/m_q respectively. The possible values for $\mu_i(S)$ are then found by elementary number theory.

THEOREM 14.1. *At an unramified prime divisor p every local algebra S_p split by K_P/k_p is a component([34]) $(S)_p$ of an algebra S split by K/k, while at a ramified prime divisor $p=p_i$ of p-degree m_i $(i=1, \cdots, t)$ the index of the group $S^{(p)}$ in $\mathfrak{S}^{(p)}$ is*

$$[\mathfrak{S}^{(p)}:S^{(p)}] = m_i[(m_i, \mathfrak{n}_i)]^{-1}, \tag{1}$$

$$\mathfrak{n}_i = \text{l.c.m.}\ [m_1, \cdots, m_{i-1}, m_{i+1}, \cdots, m_t, J(\Gamma)].$$

Specifically, a local algebra $S^{(p)}$ is a component $(S)_p$ of some S if and only if its invariant satisfies the condition

$$(m_i, \mathfrak{n}_i)\mu(S^{(p)}) \equiv 0 \pmod{n}.$$

That the indices (1) are drastically limited is indicated by the following calculation. In $[\mathfrak{S}:S]$ introduce as an intermediate group the composite $S\mathfrak{S}'$ and apply the reduction principle (2) of §9. Then

$$[\mathfrak{S}:S] = [\mathfrak{S}:\mathfrak{S}'S][\mathfrak{S}'S:S] = [\mathfrak{S}:\mathfrak{S}'S][\mathfrak{S}':S']. \tag{2}$$

The second index was computed in Theorem 4.1 to be $J(\Gamma)$. The first index may be transformed by the homomorphism which carries each ideal algebra $\mathfrak{S}$ into the vector $\{(\mathfrak{S})_1, \cdots, (\mathfrak{S})_t\}$ of its components $(\mathfrak{S})_i$ at the ramified prime divisors p_i. Under this homomorphism each algebra of the join $\mathfrak{S}'S$ goes into the vector of components $\{(S)_1, \cdots, (S)_t\}$ of the actual algebra S. Hence, by the isomorphism principle,

$$[\mathfrak{S}:\mathfrak{S}'S] = [\{(\mathfrak{S})_1, \cdots, (\mathfrak{S})_t\}:\{(S)_1, \cdots, (S)_t\}].$$

On the right we obtain a larger group if we allow distinct $(S)_i$'s to arise from different algebras S. The group index is then that for a direct product, so that

$$[\mathfrak{S}:\mathfrak{S}'S] \geq \prod_{i=1}^{t} [\mathfrak{S}^{(p_i)}:S^{(p_i)}] = \prod_{i=1}^{t} m_i/(m_i, \mathfrak{n}_i).$$

If in (2) $[\mathfrak{S}:S]=N$ and $[\mathfrak{S}':S']=J(\Gamma)$ are evaluated as in Theorems 4.2 and 4.1, we obtain

([33]) Notation as in (*13.1).

([34]) For the time being, we denote the p-component of an actual algebra S by $(S)_p$.

THEOREM 14.2. *The local algebra indices of* (1) *are limited by*

$$\prod_{i=1}^{t} [\mathfrak{S}^{(p_i)}:S^{(p_i)}] \leqq N/J(\Gamma). \tag{3}$$

However, these local indices are not necessarily all 1.

THEOREM 14.3. *Over any algebraic number field k there is an (abelian) normal extension K for which $[\mathfrak{S}^{(p)}:S^{(p)}]>1$ for at least one prime divisor p.*

Proof. If k is not the rational number field, there is at least one[35] rational prime l which has in k a prime ideal divisor p of absolute local degree greater than 1. We assert then that $[\mathfrak{S}^{(p)}:S^{(p)}]>1$. As in Theorem 13.4, the local class field theory gives an extension A_P/k_p which is abelian of type $(l, l, \cdots, l)$ with at least three generators. According to the Grunwald existence theorem there is a field A/k with the same Galois group as A_P/k_p which has at p the local component A_P/k_p and which has at each other prime ideal divisor q of l some specified local component A_Q/k_q of degree (say) l^2. Then in the index $[\mathfrak{S}^{(p)}:S^{(p)}]$ of (1), the p-degree m_1 corresponding to $p_1=p$ is at least l^3. If any $q=q_i$ is ramified, the corresponding degree m_i is at most l^2, by construction. Finally, if there is a ramified p_j not a divisor of l, this p_j cannot have higher ramifications because the degree of the extension A/k is a power of l. By the Hilbert theory, the p_j-degree m_j can then be at most l^2. In (1), each degree m_i (with $i\neq 1$) is at most l^2, while $J(\Gamma)$ is l. It follows that $[\mathfrak{S}^{(p)}:S^{(p)}]\geqq l$.

It remains to consider the case when $k=R$, the field of rational numbers. The extension $K=R(2^{1/2}, (-1)^{1/2}, 5^{1/2})$ is abelian of type (2, 2, 2). The prime ideal (2) of R is ramified in six of the quadratic subfields of K, while in the seventh field $R(5^{1/2})$ it is inert. By the Hilbert theory it follows that $(2)=P^4$ in K. Therefore the degree m_1 corresponding to $p_1=(2)$ is $m_1=8$, any $q\neq 2$ in R has no higher ramifications, hence has degree m_q at most 4. By formula (1) it follows that[36] $[\mathfrak{S}^{(p)}:S^{(p)}]\geqq 2$.

Certain properties of the norm residue symbol[37] suggest also the consideration of algebras with unit factor sets

$\mathfrak{U}^{(p)}$ = the group of classes of algebras $\mathfrak{S}^{(p)}$ with factor sets of P-adic units.

By Theorem 12.1 the group $\mathfrak{U}^{(p)}$ has order e_p. As in Theorem 6.2 one may prove

THEOREM 14.4. *An algebra S has its p-component in $\mathfrak{U}^{(p)}$ if and only if S has a crossed product representation with a factor set F of numbers relatively prime to p.*

---

(35) For example, one might select any l which is ramified in k/R.

(36) The extension K might also have been constructed by local class field theory. Indeed, 2, -1, and 5 are the generators of the factor group R_2^*/R_2^{*2} in the dyadic field R_2.

(37) See Chapter V, §27, Theorem 27.4.

The previous methods also enable us to determine when a local algebra $\mathfrak{U}^{(p)}$ with an invariant $x[n/e_p]$ (mod n) can be realized as the component of an actual algebra S.

THEOREM 14.5. *At a ramified prime divisor $p=p_i$ the number of algebras $\mathfrak{U}_i=\mathfrak{U}^{(p_i)}$ which cannot be realized as components $S_i=S^{(p_i)}$ of an actual algebra S is given by*

$$[\mathfrak{U}_i:\mathfrak{U}_i\cap(S)_i] = e_i/(e_i, \bar{n}_i), \tag{4}$$

where $\bar{n}_i$ is determined(38) *as in* (1); $i=1, 2, \cdots, t$.

By the devices used in the proof of Theorem 14.3 one may demonstrate that this index necessarily exceeds 1 if p is a prime divisor of k of absolute degree at least 3, and if K is suitably constructed.

15. **An invariant of a group pair.** For the convenient statement of subsequent results we need a certain *"deficiency"* invariant for homomorphisms of groups. Consider any abelian group R for which Γ is a group of operators, with a Γ-allowable subgroup S(39), and

(1) $\mathfrak{R} = R/S$, a group with operators in Γ,

(2) ϕ, the natural homomorphism of R on $R/S = R\phi$,

(3) $\mathfrak{E}(R) = G$, a group extension of R by Γ.

Since S is a Γ-allowable subgroup of R, it follows that S is a normal subgroup of the extension G, so that the quotient group G/S is a group extension of R/S by Γ. The set of all such extensions $G\phi=G/S$ forms a subgroup of the group of all extensions of $\mathfrak{R}$ [6, 43]:

(4) $\mathfrak{E}(\mathfrak{R}) = \mathfrak{G}$, a group extension of $\mathfrak{R}$ by Γ.

The index of this subgroup in G is our deficiency index.

DEFINITION. *If the homomorphism ϕ of R has S as the subgroup of elements mapped into* 1, *the deficiency of R modulo S is*

(*15.1) $\omega(R, S)=[\mathfrak{E}(R\phi):\{\mathfrak{E}(R)\}\phi]$.

The special extensions $G\phi$ may also be described in the following terms, due essentially to A. H. Clifford. Given an extension G of R, an extension $\mathfrak{G}$ of $\mathfrak{R}$ by the same group Γ, and a homomorphism ϕ of R to $\mathfrak{R}$, we say that G is a ϕ-prolongation of $\mathfrak{G}$ by S if and only if the homomorphism ϕ can be so extended to a homomorphism of G to $\mathfrak{G}$ that the subgroup of those elements of G mapped into 1 is exactly the given subgroup S of R. The extensions $G\phi$

(38) For the notation see the preceding developments.

(39) That is; s^σ is in S for every s in S and σ in Γ.

of our definition can then be characterized as those extensions $\mathfrak{G}$ of $\mathfrak{R}$ which admit a ϕ-prolongation by S. The problem of prolonging a given group $\mathfrak{G}$ by S may also be described as the problem of superimposing two group pairs,

$$\mathfrak{G} \supset \mathfrak{R} \supset 1, \qquad R \supset S \supset 1, \qquad R/S = \mathfrak{R},$$

in such a fashion that they "overlap" properly in a combined group

$$G \supset R \supset S \supset 1, \qquad G/S = \mathfrak{G} \text{ extending } R/S = \mathfrak{R}.$$

This combined group may be viewed as a specially restricted extension of S by $\mathfrak{G}$, where S is considered as a group with operators in $\mathfrak{G}$ according to the natural rule $S^{u_\sigma} = S^\sigma$ for all elements u_σ in a coset σ of $\mathfrak{G}/\mathfrak{R}$.

In less invariant fashion, the same deficiency index may be described in terms of factor sets. If $R_{\sigma,\tau}$ is a factor set FR, then the elements $(R_{\sigma,\tau})\phi$ constitute a factor set for $R\phi = \mathfrak{R}$.

THEOREM 15.1. *The deficiency index, in terms of factor sets, is*

$$\omega(R, S) = [F(R\phi) : (FR)\phi], \tag{5}$$

provided one of these indices is finite.

Proof. Each factor set $F(R\phi)$ determines in the usual manner a group extension $\mathfrak{G} = (R\phi, \Gamma, F(R\phi))$, while a factor set $(FR)\phi$ yields under this map one of the "prolongable" extensions $G\phi = (R, \Gamma, FR)\phi$. We shall have proven (5) in accordance with the isomorphism principle of group theory, provided we show conversely that

$$(R\phi, \Gamma, F(R\phi)) = G\phi \quad \text{implies} \quad F(R\phi) = (FR)\phi.$$

Now G is given by a factor set as $G = (R, \Gamma, FR)$, so $G\phi$ also has the factor set $(FR)\phi$. The sets $F(R\phi)$, $(FR)\phi$ thus determine the same extension of $\mathfrak{R}$, hence are similar,

$$F(R\phi) = (FR)\phi[\mathfrak{R}_\sigma \mathfrak{R}_\tau^\sigma \mathfrak{R}_{\sigma\tau}^{-1}].$$

The transformation set $T\mathfrak{R}_\sigma$ can clearly be written as the homomorphic image of a transformation set TR_σ, where R_σ is chosen in the coset $\mathfrak{R}_\sigma$. Hence

$$F(R\phi) = \{(FR)[R_\sigma R_\tau^\sigma R_{\sigma\tau}^{-1}]\}\phi,$$

as required for our proof.

16. **Factor sets of principal ideals.** In the formulas of §11, there appears the index $[F\mathfrak{A}' : (F')]$, known to be finite. Between these two groups we may insert the group $F(A')$ of factor sets of principal ideals. This group need not be the same as the group $(F') = (FA')$ of principal ideals derived from factor sets of numbers. In fact, the index of (F') in $F(A')$ is exactly a deficiency index, in the sense of §15, for the homomorphism $A' \to (A')$ belonging to the subgroup E of units; or

$$[F(A'):(F')] = [F(A'):(FA')] = \omega(A', E), \tag{1}$$

$$[F\mathfrak{A}':(F')] = [F\mathfrak{A}':F(A')]\omega(A', E). \tag{2}$$

The first index here suggests the use of the group of ideal classes in K. Specifically, set

(*16.1) $\mathfrak{A}^\dagger$ = the ideal class determined by the ideal $\mathfrak{A}$.

Then $\mathfrak{A}^\dagger$ will also, as usual, denote the (finite) group of all ideal classes in K. Applying to part of (2) the homomorphism $\mathfrak{A}\to\mathfrak{A}^\dagger$, we have

$$[F\mathfrak{A}':F(A')] = [(F\mathfrak{A}')^\dagger:1] = [F\mathfrak{A}'^\dagger:1][F\mathfrak{A}'^\dagger:(F\mathfrak{A}')^\dagger]^{-1}. \tag{3}$$

The second index is by definition a deficiency index

$$\omega(\mathfrak{A}', (A')) = [F\mathfrak{A}'^\dagger:(F\mathfrak{A}')^\dagger] = [F\mathfrak{A}^\dagger:(F\mathfrak{A}')^\dagger], \tag{4}$$

for $\mathfrak{A}^\dagger = \mathfrak{A}'^\dagger$. The first index $[F\mathfrak{A}'^\dagger:1]$ may be reduced to group extensions,

$$[F\mathfrak{A}^\dagger:1] = [F\mathfrak{A}^\dagger:T\mathfrak{A}^\dagger][T\mathfrak{A}^\dagger:1]. \tag{5}$$

The homomorphism which carries a vector of classes $\mathfrak{A}_\nu^\dagger$ into a transformation set $T\mathfrak{A}^\dagger$ will introduce the group of crossed characters $U\mathfrak{A}^\dagger$, as

$$\begin{aligned}[T\mathfrak{A}^\dagger:1] &= [\mathfrak{A}_\nu^\dagger:1][U\mathfrak{A}^\dagger:1]^{-1} \\ &= [\mathfrak{A}^\dagger:1]^n[U\mathfrak{A}^\dagger:\mathfrak{A}^{\dagger 1-\sigma}]^{-1}[\mathfrak{A}^{\dagger 1-\sigma}:1]^{-1}.\end{aligned}$$

The index $[\mathfrak{A}^\dagger:1]$ is the class number H, while the homomorphism $\mathfrak{A}^\dagger\to\mathfrak{A}^{\dagger 1-\sigma}$ will give

$$[T\mathfrak{A}^\dagger:1] = H^{n-1}[U\mathfrak{A}^\dagger:\mathfrak{A}^{\dagger 1-\sigma}]h. \tag{6}$$

The results (2) through (6) may now be combined as

THEOREM 16.1. *In terms of the group $\mathfrak{A}^\dagger$ of ideal classes in K and the deficiency index ω of* §15, *we may write*

$$\begin{aligned}[F\mathfrak{A}':(F')] &= H^{n-1}h\omega(A', E)[\omega(\mathfrak{A}', (A'))]^{-1}[J(\mathfrak{A}^\dagger)]^{-1}, \\ J(\mathfrak{A}^\dagger) &= [U\mathfrak{A}^\dagger:\mathfrak{A}^{\dagger 1-\sigma}][F\mathfrak{A}^\dagger:T\mathfrak{A}^\dagger]^{-1}.\end{aligned} \tag{7}$$

In this formula there appear two indices ω depending on the module M. We inquire now how each depends on the nature of M.

THEOREM 16.2. *The index $\omega(\mathfrak{A}', (A'))$ is independent of the choice of the module M, provided only that this module involves all ramified prime divisors for K/k.*

Proof. We may consider two such modules M_1 and M with $M_1 | M$. The respective indices are (4) and

$$\omega(\mathfrak{A}_1', (A_1')) = [F\mathfrak{A}_1' : (F\mathfrak{A}_1')^f],$$

where $\mathfrak{A}_1'$ (A_1') are ideals (elements) relatively prime to M_1. Hence the conclusion of the theorem will follow if we prove $(F\mathfrak{A}_1')^f = (F\mathfrak{A}')^f$. This will be the case provided the factor set $F\mathfrak{A}_1'$ can be turned into a factor set relatively prime to M by multiplication with a suitable factor set of principal ideals, as asserted in the following lemma:

LEMMA 1. *If p is unramified in K/k and $\mathfrak{F}$ is a p-primary factor set of ideals in K, there exists another factor set $\mathfrak{F}_1$ relatively prime to any specified module such that $\mathfrak{F}\mathfrak{F}_1$ is a factor set of principal ideals.*

Proof. Our chief concern is to show that $\mathfrak{F}_1$ can be chosen so as to be a factor set. Select P, a prime factor of p with the decomposition group $\Delta(P)$, and let L be the corresponding decomposition field in K/k. Since P is unramified the ideal P is present(40) in L. There is then in L an ideal $\mathfrak{C}$ relatively prime to M such that $P\mathfrak{C}$ is principal. Furthermore $\mathfrak{C}$, like all the elements of L, is invariant under all automorphisms δ of the group $\Delta(P)$. Because of this invariance, we can, in an unambiguous fashion, set up the following extension of the correspondence $\Psi(P) = \mathfrak{C}$ to all p-primary ideals (see §5). If $\mathfrak{A} = \prod_P P^{a_P e_P}$, let

$$\Psi(\mathfrak{A}) = \prod_P \mathfrak{C}^{a_P e_P}.$$

For arbitrary p-primary ideals $\mathfrak{A}$, $\mathfrak{B}$ this function has, according to its origin, the properties

(8) $\Psi(\mathfrak{A})$ is relatively prime to M, $\mathfrak{A}\Psi(\mathfrak{A})$ is principal,

(9) $\Psi(\mathfrak{A}\mathfrak{B}) = \Psi(\mathfrak{A})\Psi(\mathfrak{B}), \qquad \Psi(\mathfrak{B}^\tau) = [\Psi(\mathfrak{B})]^\tau,$

for every τ in Γ. In other words, the mapping $\mathfrak{A} \to \Psi(\mathfrak{A})$ is an operator homomorphism of the group $\mathfrak{A}^{(p)}$ of p-primary ideals on a subgroup of the group $\mathfrak{A}'$. It follows that the map $\Psi(\mathfrak{F})$ of a p-primary factor set will be itself a factor set. Consequently, the product $\mathfrak{F}\Psi(\mathfrak{F})$ will be a factor set of principal ideals, as asserted in the lemma.

The dependence of the other deficiency index of (7) on M is exhibited by

(10) $M \mid M_1$ implies $\omega(A_1', E) \leqq \omega(A', E)$.

For, the index $\omega(A', E)$ of (1) involves the subgroup (FA'), which may be written as the intersection $(FA') = (FA) \cap F(A')$, since a factor set of principal ideals is relatively prime to M if and only if the generators of the ideals are relatively prime to M. Hence

(40) Subject to the usual convention that an ideal in L is to be identified with its extended ideal in K.

$$\omega(A', E) = [F(A'):(FA) \cap F(A')] = [(FA)F(A'):(FA)]. \tag{11}$$

In this form the conclusion (10) is immediate. It implies that $\omega(A', E)$ is eventually independent of the module M. In order to obtain a more precise result we must employ a roundabout method.

THEOREM 16.3. *The index $\omega(A', E)$ is independent of the module M, provided only that this module M involves all ramified prime divisors of K/k.*

Proof. If the formula (11) is substituted in the final formula (5) of Theorem 11.2 for the basic index J, one obtains a relation (see (15) below) between various indices which are all independent of the module M, with the exception of the indices $\omega(A', E)$, $\omega(\mathfrak{A}', (A'))$, and $[F':F'']$. The latter index was shown in Theorem 13.1 to be independent of M, while we have just shown that the index $\omega(\mathfrak{A}', (A'))$ is independent of a module M of the type under consideration. This gives the result of Theorem 16.3.

THEOREM 16.4. *If K/k is a normal extension such that K and all subfields L over which K is cyclic have class number 1, then $\omega(A', E) = 1$.*

Proof. The proof of the theorem depends on a choice of special prime elements in K. Let p be an unramified prime ideal for K/k, $\mathfrak{A}^{(p)}$ the group of p-primary ideals in K, P a typical prime divisor of p in K, $\Delta = \{\delta\}$ the associated decomposition group. Then P is an ideal of the decomposition field L_Δ, over which K is a cyclic extension. By hypothesis L_Δ has class number 1, hence

$$P = (\pi), \qquad \text{with } \pi \text{ in } L_\Delta \text{ and } \pi^\delta = \pi. \tag{12}$$

Next write Γ in terms of cosets

$$\Gamma = \sigma_1\Delta + \sigma_2\Delta + \cdots + \sigma_g\Delta, \qquad g = [\Gamma:\Delta].$$

The prime ideal factorization of p in K is given as

$$p = \prod_{i=1}^{g} P^{\sigma_i}, \qquad P^{\sigma_i\delta} = (P^\delta)^{\sigma_i} = P^{\sigma_i}. \tag{13}$$

Now select for each prime ideal $P^{\sigma_i} = P_i$ the prime element $\pi^{\sigma_i} = \pi_i$ as generator, $P_i = (\pi_i)$ in K. Let $\Pi^{(p)}$ be the group of power products of all the prime elements π_i. By construction $\mathfrak{A}^{(p)}$ and $\Pi^{(p)}$ are groups with Γ as operator group. We assert that they are operator isomorphic with respect to Γ. For the proof we use the correspondence

$$\pi_1^{a_1} \cdots \pi_g^{a_g} \to (\pi_1^{a_1} \cdots \pi_g^{a_g}) = P^{a_1\sigma_1} \cdots P^{a_g\sigma_g}. \tag{14}$$

We have to investigate the effect of an element σ in Γ. By the coset decomposition of Γ, we get $\sigma = \sigma_i\delta$, $\sigma\sigma_j = \sigma_{j'}\delta_j$, where δ_j in Δ and j' in $\{1, \cdots, g\}$ are both determined by j. Consequently

$$\left(\prod_{j=1}^{g}\pi_j^{e_j a_j}\right)^{\sigma} = \prod_{j=1}^{g}\pi^{\sigma e_j a_j} = \prod_{j=1}^{g}\pi^{e_{j'} b_j a_j} = \prod_{j=1}^{g}\pi^{e_{j'} a_j}$$

is an element of $\Pi^{(p)}$. Similarly

$$\left(\prod_{j=1}^{g}P^{e_j a_j}\right)^{\sigma} = \prod_{j=1}^{g}P^{\sigma e_j a_j} = \prod_{j=1}^{g}P^{e_{j'} b_j a_j} = \prod_{j=1}^{g}P^{e_{j'} a_j}.$$

Thus the correspondence (14) is an operator isomorphism with Γ as allowed set of operators.

Choose now for every p relatively prime to M such a group $\Pi^{(p)}$, and let B' be the group of numbers generated by all groups $\Pi^{(p)}$. Then $\mathfrak{A}'$ is operator isomorphic with B' under the natural extension of the mapping (14). For the proof remark that $\mathfrak{A}'$ is the direct product[41] of the groups $\mathfrak{A}^{(p)}$. By the correspondence $B' \leftrightarrow (B')$ we map B' in an operator isomorphic fashion in $\mathfrak{A}'$. This rule is, in fact, the rule (14). It is an isomorphism by the behaviour of the exponents. Finally, we remark that every element A in A' has a unique representation $A = EB$ where E is a unit of K and B an element of B'. By the prime ideal decomposition of (A) we first determine an element B in B' such that $(B) = (A)$. Then $(AB^{-1}) = 1$, whence $A = BE$ with a unit. This representation is unique, for $A = E_1B_1 = EB$ implies $(B) = (B_1)$ and therefore $B = B_1$ and $E = E_1$. Consequently, A' is equal to the direct product $B' \times E$ with Γ as operator group. Therefore by Lemma 2 of §9[42], $\omega(A', E) = 1$.

To summarize the results thus far obtained, use the basic formula (5) of §11 with the value for $[F\mathfrak{A}' : (F')]$ obtained in Theorem 16.1 and the formula for $[F' : F'']$ obtained in §13. The latter formula involves the product $\prod e_p$ of the ramification orders for all prime divisors p. Part of this will cancel against the product $\prod e_p$ over all finite prime divisors, as in (5) of §11. There remains the product $\prod e_{p_\infty}$ over the infinite p's. This is just 2^ρ, where ρ is the number of infinite prime divisors of k ramified in K/k. All told, the basic index $J = J(\Gamma)$ is

$$J = n^* n^{-1}\omega(A', E)\omega(\mathfrak{A}', (A'))^{-1}J(E)\left[J(\mathfrak{A}^\delta)\right]^{-1}2^\rho,$$

$$J(E) = [UE : E^{1-\sigma}][FE : TE]^{-1}, \tag{15}$$

$$J(\mathfrak{A}^\delta) = [U\mathfrak{A}^\delta : \mathfrak{A}^{\delta 1-\sigma}][F\mathfrak{A}^\delta : T\mathfrak{A}^\delta]^{-1}.$$

Chapter III. Factor-sets of units

17. The Herbrand unit group. The analysis of the index $J(E)$ entering in our formula is based on the explicit structure of the group E of units. This in turn depends on the classification of the infinite prime divisors of k and K. Recall that a prime divisor $p = p_\infty$ of k is called real or complex according as

[41] With Γ as set of operators.

[42] Lemma 2 of §9 obviously generalizes to the kind of direct product we consider here.

the corresponding complete field k_p is the real or the complex field (see §2). The infinite prime divisors of k then fall into three classes according to the types of their divisors in K (see [5]):

(1) $p_{\infty,i}$ $(i=1, \cdots, \rho_1)$ real in k, with n real divisors in K,

(2) $p_{\infty,i}$ $(i=\rho_1+1, \cdots, r_1)$ real in k, with $n/2$ complex divisors in K,

(3) $p_{\infty,i}$ $(i=r_1+1, \cdots, r_1+r_2)$ complex in k, with n complex divisors in K.

In each case let $P_{\infty,i}$ be any one prime divisor of $p_{\infty,i}$, while $\Delta(P_{\infty,i})$ is the corresponding decomposition group (composed of all automorphisms of Γ which leave fixed the valuation belonging to $P_{\infty,i}$). For a prime divisor of the first or the third type, $\Delta(P_{\infty,i})$ consists only of the identity, and $P_\infty \mid p_\infty$ is said to be *unramified*. For a prime divisor of the second type, $\Delta(P_{\infty,i})$ is a cyclic group of order 2, while $P_\infty \mid p_\infty$ is *ramified*. For all cases, set

(4) δ_i = the generator of the decomposition group $\Delta(P_{\infty,i})$,

(5) K_i = the decomposition field of $P_{\infty,i}$ (the field left invariant by δ_i).

Then in case (2), δ_i has order $2=[K:K_i]$, while in cases (1) and (3), $\delta_i=1$ and $K=K_i$.

The unit group of k contains $r=r_1+r_2-1$ independent units $\epsilon_1, \cdots, \epsilon_r$ which, together with a suitable root of unity, generate the whole group of units in k. For the same reason there will be a maximum of $R=n\rho_1+(n/2)(r_1-\rho_1)+nr_2-1$ independent units in K. Herbrand's generalization of Minkowski's theorem [5, 23, 25, 26] asserts that one can find in K $r+1$ units $H_1, H_2, \cdots, H_{r+1}$ which with their conjugates H_i^σ and with the units $\epsilon_1, \epsilon_2, \cdots, \epsilon_r$ will generate a group of units which has finite index in the whole group E of all units in K. Furthermore, the multiplicative relations between these units are all consequences of the simple relations[43]

$$H_i^{\delta_i} = H_i, \qquad N_i H_i = 1 \qquad (i = 1, \cdots, r+1), \tag{6}$$

where N_i denotes the relative norm of K_i/k. By an Herbrand group E^* of units in K we mean a group generated by such a basis $\epsilon_i, H_i, H_i^\sigma, \cdots$. In other words, E^* is a direct product

$$E^* = \epsilon_1 \times \cdots \times \epsilon_r \times H_1 \times \cdots \times H_{r+1} \tag{7}$$

of infinite cyclic groups ϵ_i, each generated by a basic unit ϵ_i of k, and of groups H_i, where each H_i is generated by a unit H_i and its conjugates

$$H_i = \{H_i, H_i^\sigma, \cdots\}. \tag{8}$$

[43] This provides the proper number of units. If $p_{\infty,i}$ is unramified, H_i and its conjugates provide $n-1$ independent units, ϵ_i one more; if $p_{\infty,i}$ is ramified, $H_i=P_i$ and its conjugates H_i^σ provide $n/2-1$ independent units, the total number being in agreement with the value of R.

The groups ϵ_i and H_i which figure in (7) are groups which have Γ as group of operators. They can be described in purely group theoretical language. Thus ϵ_i is an infinite cyclic group with a generator ϵ_i which is invariant under all operators of Γ.

To describe the group H_i belonging to an unramified prime divisor $p_{\infty,i}$, consider first the operator free abelian group $G(\Gamma)$ belonging to any finite group Γ of order n. As in Hall's paper ([17]) the group $G(\Gamma)$ is an abelian group with n free generators G^σ, one for each element σ in Γ. To these generators the elements τ of Γ are to be applied by the rule $(G^\sigma)^\tau = G^{\tau\sigma}$. In this group $G(\Gamma)$ the elements left unaltered by all operators of Γ form an infinite cyclic subgroup $N(\Gamma)$ which is generated by the "norm" $NG = G^{1+\sigma+\tau+\cdots}$. The quotient group $G(\Gamma)/N(\Gamma)$ belonging to this subgroup $N(\Gamma)$ is then isomorphic to the unit group $H_i = H$. This quotient group is an abelian group with n generators H^σ(44) subject only to the relation $NH = 1$. We call this quotient group $H = G(\Gamma)/N(\Gamma)$ an *abstract unramified unit group*:

$$(9) \qquad H = \{H, H^\sigma, H^\tau, \cdots\}, \qquad \prod_\sigma H^\sigma = 1, \qquad (H^\sigma)^\tau = H^{\tau\sigma}.$$

In the ramified case let $\delta = \delta_i$ be the generator of the decomposition group of some ramified prime divisor $P_{\infty,i}$, and let Γ be decomposed modulo the subgroup $\{1, \delta\}$ into left cosets

$$\Gamma = \sigma_1\{1, \delta\} + \sigma_2\{1, \delta\} + \cdots + \sigma_g\{1, \delta\}, \qquad g = n/2.$$

Then the *abstract δ-ramified unit group* is an abelian group described by generators

$$(10) \qquad \mathfrak{P}^{\sigma_1}, \mathfrak{P}^{\sigma_2}, \cdots, \mathfrak{P}^{\sigma_g}, \qquad \text{with } \mathfrak{P}^{\sigma_1}\mathfrak{P}^{\sigma_2}\cdots\mathfrak{P}^{\sigma_g} = 1.$$

The operators of Γ are applied to this group according to the rules

$$(11) \qquad (\mathfrak{P}^{\sigma i})^\tau = \mathfrak{P}^{\tau\sigma i}, \qquad \mathfrak{P}^{\sigma i\delta} = \mathfrak{P}^{\sigma i}, \qquad i = 1, 2, \cdots, g.$$

The Herbrand group E^* of (7) is thus represented as a direct product whose factors are

r infinite cyclic groups,

$\rho_1 + r_2$ abstract unramified unit groups,

$r_1 - \rho_1$ abstract δ_i-ramified unit groups belonging to the prime divisors $p_{\infty,i}$ of type (2).

18. Reduction to the Herbrand unit group. The formula of Theorem 11.2 involves the units of K in the index

(44) Each H^σ is the coset of the original generator G^σ.

$$J(E) = [UE:E^{1-\sigma}][FE:TE]^{-1}. \tag{1}$$

We desire to relate this index to the corresponding index for the Herbrand unit group E^*

$$J(E^*) = [UE^*:E^{*1-\sigma}][FE^*:TE^*]^{-1}. \tag{2}$$

THEOREM 18.1. *The indices involved in $J(E^*)$ are both finite, and*

$$J(E) = J(E^*)\{[E:E^*]^{n-1}[\epsilon:\epsilon \cap E^*][FE:FE^*]^{-1}\} \tag{3}$$

where all the indices on the right are finite, and where $\epsilon \cap E^*$ is the group of all units of k which lie in the Herbrand group E^*.

Proof. We first show why the indices in (3) are finite. If E_σ is the group of all vectors with components in the group E, with a similar meaning for E_σ^*, then E_σ is the n-fold direct product of E by itself, so that

$$[E_\sigma:E_\sigma^*] = [E:E^*]^n$$

is finite. Apply next the homomorphism $E_\sigma \to TE_\sigma$. By definition, the subgroup of E_σ mapped onto the identity is the group UE of crossed characters. Then, by the homomorphism principle of §9,

$$[E:E^*]^n = [TE:TE^*][UE:UE^*],$$

where both indices on the right must be finite. This may be written as

$$[UE:UE^*] = [E:E^*]^n[TE:TE^*]^{-1}. \tag{4}$$

Secondly, apply the homomorphism which carries a unit E into a vector with components $E^{1-\sigma}$. The subgroup mapped on the identity is ϵ, the group of units in k. Hence

$$[E:E^*] = [E^{1-\sigma}:E^{*1-\sigma}][\epsilon:\epsilon \cap E^*], \tag{5}$$

where both indices on the right are finite.

Next we return to the index $J(E)$. In the numerator the index $[UE:E^{1-\sigma}]$ suggests two chains of groups $UE \supset E^{1-\sigma} \supset E^{*1-\sigma}$ and $UE \supset UE^* \supset E^{*1-\sigma}$ which join the same groups UE and $E^{*1-\sigma}$. The whole index $[UE:E^{*1-\sigma}]$ is known to be finite. It can be computed in two ways from these chains, with the result

$$[UE:E^{1-\sigma}] = [UE^*:E^{*1-\sigma}][UE:UE^*][E^{1-\sigma}:E^{*1-\sigma}]^{-1}.$$

Inserting the values of these indices from (4) and (5), we get

$$[UE:E^{1-\sigma}] = [UE^*:E^{*1-\sigma}][E:E^*]^{n-1}[\epsilon:\epsilon \cap E^*][TE:TE^*]^{-1}. \tag{6}$$

One may similarly run two chains from FE to TE^*, through the intermediate

groups FE^* and TE. The indices $[FE:TE]$, $[TE:TE^*]$ are known to be finite, hence the indices are all finite and

$$[FE:TE] = [FE^*:TE^*][FE:FE^*][TE:TE^*]^{-1}. \tag{7}$$

Division of formula (6) by (7) yields the desired reduction formula (3).

19. **The correction factor for the reduction.** The replacement of the index $J(E)$ by the corresponding index $J(E^*)$ for the Herbrand unit group involves a "correction factor" found in equation (3) of §18,

$$\phi = J(E)/J(E^*) = [E:E^*]^{n-1}[\epsilon:\epsilon \cap E^*][FE:FE^*]^{-1}. \tag{1}$$

In the computations for the cyclic case in the classical theory it turns out([45]) that the corresponding correction factor is always 1, which is to say that the indices involving the units can be computed directly from the Herbrand group E^* itself instead of from E. It would be convenient to have also $\phi=1$ in general([46]). This is not the case. We shall prove $\phi=1/2$ for a suitably selected pair of fields K, k.

Before actually proving this assertion we shall establish some lemmas reducing the problem to a simpler problem involving "normalized" factor sets.

Let Γ be the four group, with generators γ, β:

$$\Gamma = \{1, \gamma, \beta, \gamma\beta\}, \qquad \gamma^2 = \beta^2 = 1, \qquad \beta\gamma = \gamma\beta = \delta. \tag{2}$$

A factor set $E_{\sigma,\tau}$ determines a group extension of the unit group E by the group Γ. Any such group extension is completely determined by the multiplication table for elements u_γ and u_β which represent the two generating cosets γ and β in the extension. This multiplication table([47]) has the following form

$$u_\gamma E = E^\gamma u_\gamma, \qquad u_\beta E = E^\beta u_\beta, \qquad \text{(for any } E\text{)}, \tag{3}$$

$$u_\gamma^2 = C, \qquad u_\beta^2 = B, \qquad u_\beta u_\gamma = D u_\gamma u_\beta, \tag{4}$$

where C, B, and D lie in the group E. The multiplication determined by this table will be associative if and only if C, B, and D satisfy the conditions

$$C^\gamma = C, \qquad B^\beta = B, \qquad D^{1+\gamma} = C^{\beta-1}, \qquad D^{1+\beta} = B^{1-\gamma}. \tag{5}$$

The necessity of these conditions is readily proved by "multiplying out" in two different associations the products u_γ^3, u_β^3, $u_\beta^2 u_\gamma$ and $u_\beta u_\gamma^2$. Thus, any group

---

([45]) See Theorem 19.2 below.

([46]) This was actually conjectured by Deuring in a letter to Hasse in 1933, relating to some investigations on the principal ideal theorem of Hilbert.

([47]) Similar normalizations considered as generalizations of the formulas for cyclic algebras were used in the investigations of Dickson on crossed products [14]. Our normalization is a special case of the normalization of factor sets for abelian groups as set forth in [43, pp. 95–97].

extension of E by Γ is determined by a vector of three constants $\{C, B, D\}$ satisfying (5).

The fact that any group extension can be normalized in this form means that there is a homomorphic mapping of the factor sets $E_{\sigma,\tau}$ onto the "normalized" factor sets. To obtain explicitly this homomorphism, let the factor set $E_{\sigma,\tau}$ generate a group extension in which the representatives v_σ of the cosets σ multiply according to the table $v_\sigma v_\tau = E_{\sigma,\tau} v_{\sigma\tau}$. A corresponding normalized factor set then can be obtained by setting $u_\gamma = v_\gamma$, $u_\beta = v_\beta$ and computing C, B, and D from the table of the $E_{\sigma,\tau}$'s. One finds

$$C = E_{\gamma,\gamma}E_{1,1}, \qquad B = E_{\beta,\beta}E_{1,1}, \qquad D = E_{\beta,\gamma}E_{\gamma,\beta}^{-1}. \tag{6}$$

These formulas provide a homomorphic mapping

$$FE = E_{\sigma,\tau} \to \{C, B, D\}. \tag{7}$$

What factor sets $E^{(0)}_{\sigma,\tau}$ are mapped by this homomorphism onto the identity? If $E^{(0)}_{\sigma,\tau}$ belongs to the multiplication table $v_\sigma v_\tau = E^{(0)}_{\sigma,\tau} v_{\sigma\tau}$, then $u_\beta = v_\beta$ and $u_\gamma = v_\gamma$ multiply as in (4) with constants $C = B = D = 1$. Furthermore the two remaining coset representatives must have the form $v_1 = E_1$, $v_\delta = E_\delta v_\beta v_\gamma$, where E_1 and E_δ are elements([48]) of E. From these constants we may calculate the whole multiplication table, getting

$$E^{(0)}_{\sigma,\tau} = E_\sigma E^{\sigma}_{\tau} / E_{\sigma\tau}, \qquad \text{where } E_\beta = E_\gamma = 1. \tag{8}$$

Furthermore, the given factor set $E^{(0)}_{\sigma,\tau}$ uniquely determines the two constants E_1 and E_δ as $E_1 = E^{(0)}_{1,1}$, $E_\delta = E^{(0)}_{\delta,\beta}$. Conversely, any two elements E_1 and E_δ determine through the formulas (8) a factor set $E^{(0)}_{\sigma,\tau}$ which has the property that it is mapped by the homomorphism (7) on the identity. Hence (8) provides an isomorphism

$$\{E_1, E_\delta\} \leftrightarrow E^{(0)}_{\sigma,\tau} \tag{9}$$

between the group of pairs E_1, E_δ and the group $E^{(0)}_{\sigma,\tau}$. An analogous isomorphism applies to the group E^*.

Now apply the homomorphism (7) to the index $[FE : FE^*]$. This maps FE on the whole group of vectors $\{C, B, D\}$, because any such vector determines a group extension and hence a factor set of which it is the map. Furthermore, this homomorphism carries the group FE^* of factor sets belonging to the Herbrand subgroup E^* into the group $\{C^*, B^*, D^*\}$ of vectors from E^* satisfying (5). Therefore, by the homomorphism principle,

$$[FE : FE^*] = [\{C, B, D\} : \{C^*, B^*, D^*\}]\,[E^{(0)}_{\sigma,\tau} : (E^*_{\sigma,\tau})^{(0)}].$$

The second index, involving the group of factor sets mapped by (7) into the identity, may be computed by (9) as

([48]) That is to say, they are part of a vector E_σ used to transform a factor set.

$$[E^{(0)}_{\sigma,\tau}:(E^*_{\sigma,\tau})^{(0)}] = [\{E_1, E_3\}:\{E_1^*, E_3^*\}] = [E:E^*]^2.$$

This proves

LEMMA 1. *If Γ is the four group, the factor set index is*

$$[FE:FE^*] = [\{C, B, D\}:\{C^*, B^*, D^*\}][E:E^*]^2, \tag{10}$$

where the elements C, B, D of E (or C^*, B^*, and D^* of E^*) are restricted by the associativity conditions (5).

LEMMA 2. *If Γ is the four group, the unit "correction factor" of* (1) *is*

$$\phi = [E:E^*][D:D^*]^{-1}[\epsilon:\epsilon \cap E^*]^{-1}, \tag{11}$$

where D is the group of all elements in E for which there exist solutions B and C of the equations (5), *and D^* is analogously defined in E^*.

Proof. The group of vectors $\{C, B, D\}$ of equation (10) may be mapped on the group D described in the lemma by the homomorphism

$$\{C, B, D\} \to D. \tag{12}$$

The vectors thereby mapped into the identity ($D=1$) are the vectors $\{C, B\}$ which satisfy the relations (5) with $D=1$, or $C^\gamma = C^\beta = C$, $B^\beta = B^\gamma = B$. These relations mean that C and B are both invariant under all elements of the Galois group Γ, which is to say that C and B are both units in the group ϵ. The group mapped by (12) on the identity is therefore $\{\epsilon, \epsilon, 1\}$, so

$$[\{C, B, D\}:\{C^*, B^*, D^*\}] = [D:D^*][\epsilon:\epsilon \cap E^*]^2. \tag{13}$$

Inserting the values (13) and (10) in the correction index (1) we get the assertion of the lemma.

We remark in passing that the conclusion (11) is still valid if Γ is an abelian group which is the direct product of two cyclic groups with generators β and γ of respective orders s and t, provided the condition (5) is replaced by

$$C^\gamma = C, \quad B^\beta = B, \quad D^{1+\gamma+\cdots+\gamma^{t-1}} = C^{\beta-1}, \quad D^{1+\beta+\cdots+\beta^{s-1}} = B^{1-\gamma}. \tag{14}$$

Consider now the particular field[49] $K = R(15^{1/2}, (-2)^{1/2})$ which is normal over the field of rational numbers R and has the four group as Galois group.

LEMMA 3. *The only units in $R(15^{1/2}, (-2)^{1/2})$ are $\pm\eta^m$, where $\eta = 4+15^{1/2}$.*

Proof. One verifies at once that the only roots of unity present in K are ± 1. In the quadratic subfield $R(15^{1/2})$ it is known that the units are all of the form $\pm\eta^m$. Let N denote the norm from K to $R(15^{1/2})$. Then $N(\eta) = \eta^2$. Since

[49] In the elaborate search which ultimately led to this example, considerable use was made of the tables for units in the back of Sommer's textbook [35].

there can be only one independent unit in the totally complex field K, any other unit E in K is dependent on η, hence has $NE = \pm\eta^{m_0}$ for some $m_0 \neq 0$. We can therefore prove that $\pm\eta^m$ exhaust the whole unit group of K if we can assert that no unit E has $NE = \pm\eta$.

Since E is an integer, it is readily shown that E must have the form

$$E = (a + b15^{1/2} + c(-2)^{1/2} + d(-30)^{1/2})2^{-1} \qquad (a, b, c, d \text{ are integers}).$$

If one calculates NE, one finds that $NE = -\eta$ is impossible, while $NE = \eta$ leads to the diophantine equations

$$a^2 + 15b^2 + 2c^2 + 30d^2 = 16, \qquad 2ab + 4cd = 4.$$

By trial of a small number of cases it results that these equations have no integral solutions. Thus the lemma is proved.

THEOREM 19.1. *There exists a normal extension K/k with the four group as Galois group for which the correction factor ϕ of* (1) *is different from* 1.

Proof. Take the field $K = R(15^{1/2}, (-2)^{1/2})$. The Herbrand group E^* of this field is any group of units generated by a unit P invariant under the automorphism carrying K into its complex conjugate. The unit $P = \eta$ is such a unit, hence we may take for E^* the group of all η^m. Then E is the direct product $E = \{\pm 1\} \times E^*$, where $\{\pm 1\}$ is a cyclic group of order 2. One verifies readily that if $D = -1$ the equations (5) defining the parameter D have a solution $B = C = 1$. Hence the group D is given as $\{\pm 1\} \times D^*$. Finally, $\epsilon = \{\pm 1\}$, $\epsilon \cap E^* = 1$. With these values in the formula (11) for the correction factor we get $\phi = 2/4 = 1/2$.

The result for this particular field has been checked directly, independently of the normalized factor sets, by the laborious process of explicitly solving the associativity conditions for all factor sets $E_{\sigma,\tau}$ in the unit group $\{\pm\eta^m\}$. Using the normalization again, the unit group has been computed also for the fields $R(30^{1/2}, (-2)^{1/2})$, $R(42^{1/2}, (-2)^{1/2})$, $R(2^{1/2}, (-1)^{1/2})$. In these three cases one finds again that $\phi = 1/2$. The first two cases have unit groups isomorphic to the group of our example $R(15^{1/2}, (-2)^{1/2})$, while in the third case the unit group involves more roots of unity.

The correction factor ϕ is not the same for all fields with the four group. For the field $R(3^{1/2}, (-2)^{1/2})$ it turns out to be 1.

THEOREM 19.2. *If the Galois group Γ of K/k is cyclic, the correction factor ϕ is* 1([50]).

Proof. The proof depends on a normalization of the factor sets FE which appear in formula (1). Just as in the formula (9) of §2, each factor set of

([50]) This theorem simply asserts that the correction factor ϕ of (1) does in fact behave like the analogous correction factor of the cyclic theory.

units $E_{\sigma,\tau}$ for a cyclic group with the generator λ determines a single quantity C which functions as the usual "normalized" factor set,

$$C = \prod_{i=0}^{n-1} E_{\lambda^i,\lambda}. \tag{15}$$

The associativity condition becomes $C^\lambda = C$, so that C is a unit ϵ in the base field k. The equation (15) thus provides a homomorphic mapping $E_{\sigma,\tau} \to C$ of the group FE onto the units ϵ of k. This mapping carries FE^* into $\epsilon \cap E^*$, and one may establish as in Lemma 1 above that the factor sets $E^{(0)}_{\sigma,\tau}$ mapped by (15) into 1 have the form of transformation sets $E^{(0)}_{\sigma,\tau} = E_\sigma E^\sigma_\tau E^{-1}_{\sigma\tau}$ for which $E_\lambda = 1$. By the homomorphism principle we have

$$[FE:FE^*] = [\epsilon:\epsilon \cap E^*][E_\sigma:E_\sigma^*] = [\epsilon:\epsilon \cap E^*][E:E^*]^{n-1}.$$

According to (1) this proves $\phi = 1$.

20. **An additional reduction for unit factor sets.** In this section we give another way of reducing the basic index involving the units, $J(E) = [UE:E^{1-\sigma}]/[FE:TE]$. The result will also be used to get an estimate for $[FE:FE^*]$ and for the correction factor $J(E)/J(E^*)$. The group E may be analyzed into the subgroup ϵ and the groups

(1) $\quad \mathfrak{Z}$ = all roots of unity in K; $\zeta = \mathfrak{Z} \cap k$;

(2) $\quad B$ = all units E with NE a root of unity ζ;

(3) $\quad \Omega = B\epsilon$.

We introduce also the integer $y = [\mathfrak{Z}:1]$.

Under the homomorphism $A \to A^\nu$ the subgroup mapped on 1 is the group $\mathfrak{Z}$, hence the isomorphisms

$$E^\nu \cong E/\mathfrak{Z}, \qquad B^\nu \cong B/\mathfrak{Z}, \qquad \epsilon^\nu \cong \epsilon/\zeta. \tag{4}$$

Clearly these isomorphisms hold also if y is replaced by any proper multiple of $[\mathfrak{Z}:1]$.

THEOREM 20.1. *The index $J(E)$ is given by*

$$J(E) = J(B^\nu)J(\epsilon^\nu)J(\mathfrak{Z})\omega(E, B)\omega(B, \mathfrak{Z})/[E:B\epsilon],$$

where each ω represents a deficiency index, where for any group R $J(R) = [UR:R^{1-\sigma}]/[FR:TR]$, and where $[E:B\epsilon]$ depends on the norms within ϵ as in

$$[E:B\epsilon] = [NE:\epsilon^n]/[NE \cap \zeta:\zeta^n].$$

Since ϵ is generated by ζ and r independent units $\epsilon_1, \cdots, \epsilon_r$ of infinite order, ϵ^ν is isomorphic to a subgroup of the group generated by $\epsilon_1, \cdots, \epsilon_r$. Hence ϵ^ν is the direct product of r infinite cyclic groups. As will subsequently

be shown (§21, (4) and §23), one may then compute that $J(\epsilon^\nu) = [\Gamma:\Gamma']^{-\nu}$, where Γ' is the commutator subgroup of Γ.

The proof of Theorem 20.1 will depend on systematic use of the subgroup Ω of E, as defined in (3). The index $[E:\Omega]$ can be evaluated in several different ways.

LEMMA 1. *The index* $[E:\Omega]$ *is finite and is given by*

$$[E:\Omega] = [E^\nu:\Omega^\nu], \tag{5}$$

$$[E:\Omega] = [NE:\epsilon^n]/[NE \cap \zeta:\zeta^n], \tag{6}$$

$$[E:\Omega] = [(E^{1-\sigma})^\nu:(\Omega^{1-\sigma})^\nu][E^{1-\sigma} \cap Z_\sigma:Z^{1-\sigma}]. \tag{7}$$

Proof. If $[E:\Omega]$ is finite, the expression (5) is obtained at once by applying the homomorphism principle to the map $E \to E^\nu$, for the elements mapped thereby into 1 all lie in the subgroup Z of Ω. To get (6), apply the homomorphism $E \to NE$, with the result

$$\begin{aligned}[E:\Omega] &= [NE:N\Omega] = [NE:(NB)\epsilon^n] = [NE:(NE \cap \zeta)\epsilon^n] \\ &= [NE:\epsilon^n]/[(NE \cap \zeta)\epsilon^n:\epsilon^n],\end{aligned}$$

for $NB \subset \zeta$ by the very definition of B. Since NE is a subgroup of ϵ, $[NE:\epsilon^n]$ is less than the finite index $[\epsilon:\epsilon^n]$, which proves that $[E:\Omega]$ and all other indices here present are finite. If one observes that the intersection of $(NE \cap \zeta)$ with ϵ^n consists of all roots of unity in k which are nth powers, and hence is just ζ^n, this formula will give the result (6) by the reduction principle (§9).

Finally, to prove (7), apply the successive homomorphisms $E \to E^{1-\sigma} \to (E^{1-\sigma})^\nu$. They give

$$[E:\Omega] = [E^{1-\sigma}:\Omega^{1-\sigma}] = [(E^{1-\sigma})^\nu:(\Omega^{1-\sigma})^\nu][E^{1-\sigma} \cap Z_\sigma:\Omega^{1-\sigma} \cap Z_\sigma].$$

This will give (7), if we can prove that $\Omega^{1-\sigma} \cap Z_\sigma = Z_\sigma$. By definition, $\Omega = B\epsilon$, so $\Omega^{1-\sigma} = B^{1-\sigma}$. Since B contains all units with norms in ζ, $B \supset Z$ and $B^{1-\sigma} \cap Z_\sigma \supset Z^{1-\sigma}$. Conversely, suppose that B is a unit of norm $NB = \zeta$ with $B^{1-\sigma} = Z_\sigma$ for all σ. Then $(B^{1-\sigma})^\nu = Z_\sigma^\nu = 1$, so $(B^\nu)^{1-\sigma} = 1$, and $B^\nu = \epsilon$ is in k. Therefore $NB^\nu = \zeta^\nu = 1 = \epsilon^n$, so the unit ϵ must be a root of unity. But $B^\nu = \epsilon$, so B is also a root of unity Z. Therefore $B^{1-\sigma} \cap Z_\sigma \subset Z^{1-\sigma}$, and the lemma is established.

LEMMA 2. *If* R *and* S *are* Γ*-allowable subgroups of* K *such that* $[R:S]$ *and* $[FR:TR]$ *are both finite, then* $[FR:FS]$ *and* $[FS:TS]$ *are finite.*

Proof. The index $[FR:FS]$ can be written formally as

$$[FR:FS] = [FR:(TR)(FS)][(TR)(FS):FS].$$

The first index does not exceed $[FR:TR]$, and the second can be changed by the reduction principle to $[TR:FS \cap TR]$, which does not exceed

$[TR:TS]$. The latter index is finite, and bounded by $[R:S]^n$. Hence $[FR:FS]$ is finite. Therefore $[FS:TS]$ is finite, because it can be expressed as $[FR:TR][TR:TS]/[FR:FS]$.

The expression $J(E)$ of the theorem has as denominator the index $[FE:TE]$. This may be reduced by applying the homomorphism $A \to A^\nu$, to get

$$[FE:TE] = [(FE)^\nu:(TE)^\nu][FZ:TE \cap FZ]. \tag{8}$$

Here $[FE:TE]$ is known to be finite, while $[E:E^\nu]$ is also finite because E has a finite number of generators. Therefore $[FE:FE^\nu]$ and $[FE^\nu:TE^\nu]$ are both finite indices (Lemma 2). The latter may be introduced in (8) if we divide by the deficiency index $\omega(E, Z) = [FE^\nu:(FE)^\nu]$ (see Theorem 15.1). The subgroup Ω^ν of (3) may be introduced in terms of the finite indices $[FE^\nu:F\Omega^\nu]$ and $[TE^\nu:T\Omega^\nu]$. Then

$$\begin{aligned}[FE:TE] = {} & [FE^\nu:F\Omega^\nu][F\Omega^\nu:T\Omega^\nu][TE^\nu:T\Omega^\nu]^{-1} \\ & \times [FZ:TZ][TE \cap FZ:TZ]^{-1}[\omega(E, Z)]^{-1}.\end{aligned} \tag{9}$$

The index involving TE^ν and $T\Omega^\nu$ may be found by a homomorphic map of the group $(E^\nu)_\sigma$ of vectors with components in E^ν, as

$$\begin{aligned}[(E^\nu)_\sigma:(\Omega^\nu)_\sigma] &= [TE^\nu:T\Omega^\nu][UE^\nu:U\Omega^\nu], \\ [TE^\nu:T\Omega^\nu] &= [E^\nu:\Omega^\nu]^n[UE^\nu:U\Omega^\nu]^{-1}.\end{aligned} \tag{10}$$

In (9) we next attack $[FE^\nu:F\Omega^\nu]$ by the norm homomorphism

$$[FE^\nu:F\Omega^\nu] = [N(FE^\nu):N(F\epsilon^\nu)].$$

One proves easily that the norm of a factor set is itself a factor set. The group $N(F\epsilon^\nu) = (F\epsilon^\nu)^n$ is then contained in $F(N\epsilon^\nu) = F(\epsilon^{n\nu})$. Since each element of $\epsilon^{n\nu}$ has a unique nth root, the two groups are equal, or $N(F\epsilon^\nu) = F(\epsilon^{n\nu})$. On the other hand, the index $[F(NE^\nu):N(FE^\nu)]$ is a deficiency index $\omega(E^\nu, B^\nu)$ belonging to the norm homomorphism. All told

$$[FE^\nu:F\Omega^\nu] = [F(NE^\nu):F(\epsilon^{n\nu})]/\omega(E^\nu, B^\nu).$$

Now NE^ν and $\epsilon^{n\nu}$ are both subgroups of finite index in the group ϵ^ν, which is a direct product of r cyclic groups. Hence both NE^ν and $\epsilon^{n\nu}$ are themselves direct products of r cyclic groups. Therefore $[F(NE^\nu):T(NE^\nu)] = [F(\epsilon^{n\nu}):T(\epsilon^{n\nu})]$; in fact, they are each equal to $[\Gamma:\Gamma']^r$, by §21, (4). The factor sets on the right in the last equation may therefore be replaced by transformation quantities,

$$[FE^\nu:F\Omega^\nu] = [T(NE^\nu):T(\epsilon^{n\nu})]/\omega(E^\nu, B^\nu). \tag{11}$$

Since each NE^ν is invariant under any σ, every crossed character of Γ in the group NE^ν is an ordinary character of Γ in NE^ν. But Γ is finite and $(NE)^\nu$

contains no element of finite order, so that all these ordinary characters are 1. This means that the homomorphism $(NE^\nu) \to T(NE^\nu)_\sigma$ which carries a vector into a transformation quantity is an isomorphism. With this isomorphism, the last equation becomes

$$[FE^\nu : F\Omega^\nu] = [NE^\nu : \epsilon^{n\nu}]^n / \omega(E^\nu, B^\nu).$$

On the other hand, the homomorphism $NE \to (NE)^\nu$ gives

$$[NE^\nu : \epsilon^{n\nu}] = [NE : \epsilon^n] \cdot [NE \cap \zeta : \zeta^n]^{-1} = [E : \Omega],$$

according to (6). These two formulas, combined with (11), yield

$$[FE^\nu : F\Omega^\nu] = [E : \Omega]^n / \omega(E^\nu, B^\nu). \tag{12}$$

The deficiency index ω which is present here may be reformulated by combination with the deficiency $\omega(E, \mathfrak{Z})$ present in (9). By the norm homomorphism

$$\omega(E, \mathfrak{Z}) = [FE^\nu : (FE)^\nu] = [N(FE^\nu) : N(FE)^\nu][FB^\nu : (FB)^\nu],$$

while

$$\omega(E^\nu : B^\nu) = [F(NE^\nu) : N(FE^\nu)], \qquad \omega(E, B) = [F(NE^\nu) : N(FE)^\nu].$$

Combining these results, we find

$$\omega(E, \mathfrak{Z})\omega(E^\nu, B^\nu) = \omega(E, B)\omega(B, \mathfrak{Z}). \tag{13}$$

Turn now to the numerator $[UE : E^{1-\sigma}]$ of the expression $J(E)$. To each vector in the principal genus UE apply the homomorphism $UE \to (UE)^\nu$. Then

$$[UE : E^{1-\sigma}] = [(UE)^\nu : (E^\nu)^{1-\sigma}][U\mathfrak{Z} : E^{1-\sigma} \cap \mathfrak{Z}_\sigma]. \tag{14}$$

We may introduce the group $U(E^\nu) \supset (UE)^\nu$ by

LEMMA 3. *The index* $[UE^\nu : (UE)^\nu]$ *is finite, and is given by*

$$[UE^\nu : (UE)^\nu] = [TE \cap F\mathfrak{Z} : T\mathfrak{Z}]. \tag{15}$$

Proof. Let $U'E$ temporarily denote all those vectors E_σ of units such that the transformation set TE consists of roots of unity. Then $U'E$ contains the crossed characters UE, and the homomorphism T gives

$$[U'E : UE] = [TE \cap F\mathfrak{Z} : 1] = [TE \cap F\mathfrak{Z} : T\mathfrak{Z}][T\mathfrak{Z} : 1].$$

Since $\mathfrak{Z}$ is a finite group, the indices involved here are all finite. On the other hand, the crossed characters $U(E^\nu)$ are all νth powers of elements in this temporary group $U'E$, so the homomorphism $UE \to (UE)^\nu$ will prove

$$[U'E : UE] = [U(\Gamma^\nu) : (UE)^\nu][\mathfrak{Z}_\sigma : U\mathfrak{Z}].$$

This proves that the desired index $[U(E^v):(UE)^v]$ is finite. Furthermore we get (15) by comparing these two expressions for $[U'E:UE]$, observing that $[\mathfrak{Z}_v:U\mathfrak{Z}]=[T\mathfrak{Z}:1]$ by the very definition of the crossed characters $U\mathfrak{Z}$.

Using the result of Lemma 3, we may now rewrite (14) as

$$(16)\quad [UE:E^{1-\sigma}]=[UE^v:(E^v)^{1-\sigma}][U\mathfrak{Z}:\mathfrak{Z}^{1-\sigma}][TE\cap F\mathfrak{Z}:T\mathfrak{Z}]^{-1}[E^{1-\sigma}\cap\mathfrak{Z}_v:\mathfrak{Z}^{1-\sigma}]^{-1}.$$

The first index in this formula may be transferred to the group Ω^vS since $[E^v:\Omega^v]$ is finite, so are $[UE^v:U\Omega^v]$ and $[(E^v)^{1-\sigma}:(\Omega^v)^{1-\sigma}]$. Therefore

$$[UE^v:(E^v)^{1-\sigma}]=[UE^v:U\Omega^v][U\Omega^v:(\Omega^v)^{1-\sigma}][(E^v)^{1-\sigma}:(\Omega^v)^{1-\sigma}]^{-1}.$$

On substitution in (16) and application of (7) this gives an expression for $[UE:E^{1-\sigma}]$ as

$$(17)\quad [UE^v:U\Omega^v][U\mathfrak{Z}:\mathfrak{Z}^{1-\sigma}][U\Omega^v:(\Omega^v)^{1-\sigma}][TE\cap F\mathfrak{Z}:T\mathfrak{Z}]^{-1}[E:\Omega]^{-1}.$$

The results can now be combined to give $J(E)=[UE:E^{1-\sigma}]/[FE:TE]$ by dividing the expression (17) for the numerator by the expression (9) for the denominator, using (10) and (12) for substitutions. The result is

$$(18)\quad J(E)=J(\Omega^v)J(\mathfrak{Z})\omega(E,\mathfrak{Z})\omega(E^v,B^v)[E:\Omega]^{-1}.$$

Since $\Omega^v=B^v\epsilon^v$, and since E^v contains no roots of unity, and hence no elements in the group B^v of elements of norm 1, Ω^v is a direct product $B^v\times\epsilon^v$. The corresponding J index is therefore also a product $J(\Omega^v)=J(B^v)J(\epsilon^v)$. Putting this in (18), reducing the product of the two deficiency indices as in (13) and replacing $[E:\Omega]$ by (6), the result (18) becomes that of Theorem 1.

The indices involved in this expression for $J(E)$ are all reasonably explicit except for $J(B^v)$. As indicated in (4), the group B^v here may be replaced by B^x, where $x=[E:E^*]$ is the index of the Herbrand group. Hence the problem is essentially that of computing $[FB^x:TB^x]$, the number of group extensions of a certain subgroup B^x of the Herbrand group. The computations of §§22, 23 solve this problem for the special case when B^x is a subgroup isomorphic to the group H^* generated by all the Herbrand units of norm 1.

This type of computation can also be employed to give a new expression for the correction factor of §19.

THEOREM 20.2. *For an Herbrand subgroup E^* of E*

$$J(E)/J(E^*)=[J(B^v)/J(H^*)]J(\mathfrak{Z})\omega(E,B)\omega(B,\mathfrak{Z})[E:B\epsilon]^{-1},$$

where H^* is the subgroup of E^* consisting of all units of norm 1.

This conclusion can be obtained by a suitable direct analysis of the correction factor, using the index $x=[E:E^*]$ much as the index y of the above computation. A quicker method is to apply the computation underlying Theorem 20.1 to the group E^* as if E^* were the whole unit group of a field. Since

$\epsilon^* = E^* \cap k$, the Herbrand group is as in §17 the direct product $E^* = \epsilon^* \times H^*$, and contains no roots of unity. The result analogous to Theorem 20.1, using $y = 1$, is then

$$J(E^*) = J(H^*)J(\epsilon^*)\omega(E^*, H^*)/[E^* : H^*\epsilon^*].$$

But $E^* = H^* \times \epsilon^*$ means that the deficiency index ω here must be 1, because every factor set in H^* is automatically a factor set in E^* and thus a homomorphic image of a factor set in E^*. Furthermore ϵ^* is isomorphic to ϵ^ν, as both are free abelian groups on r generators. Hence $J(\epsilon^*) = J(\epsilon^\nu)$. This result, combined with Theorem 20.1, then gives the formula of Theorem 20.2.

21. **Extensions of unit groups.** The denominator of the index $J(E^*)$ is $[FE^* : TE^*]$, the number of group extensions of the Herbrand unit group E^* by the Galois group Γ. Since E^* is a direct product (see the end of §17), this index reduces at once to

$$[FE^* : TE^*] = \left(\prod_{i=1}^{r} [F\epsilon_i : T\epsilon_i]\right)\left(\prod_{i=1}^{r+1} [FH_i : TH_i]\right). \tag{1}$$

Here the index $[F\epsilon_i : T\epsilon_i]$ is simply the number of group extensions of an infinite cyclic group ϵ_i by the Galois group Γ, it being assumed that ϵ_i lies in the center of the resulting extension([51]). These group extensions can be counted by the character method applied in §5, to p-primary factor sets of ideals. In fact, the group extension problem is formally exactly that of finding the number of classes of p-primary factor sets in the special case when p is a prime ideal([52]) of K. The number of such classes was shown in Theorem 5.5 to be the index $[\Delta(P) : \Delta(P)']$ of the commutator group $\Delta(P)'$. Since the formal analogue of the decomposition group $\Delta(P)$ is the whole group Γ, we find in this case

$$[F\epsilon_i : T\epsilon_i] = [\Gamma : \Gamma'], \qquad \Gamma' \text{ the commutator group of } \Gamma. \tag{2}$$

The index (1) thus reduces to the following group-theoretic expression.

THEOREM 21.1. *The number of classes of associated factor sets from the Herbrand unit group E^* is*

$$[FE^* : TE^*] = [\Gamma : \Gamma']^r [FH : TH]^{n+r_0} \prod_{p_\infty} [F\mathfrak{P}_\delta : T\mathfrak{P}_\delta] \tag{3}$$

where H is the abstract unramified unit group of (9) *in* §17, *$\mathfrak{P}_\delta$ is the abstract δ-ramified unit group of* (10), (11) *in* §17, *the product is taken over all p_∞ which are ramified in K/k and $\delta = \delta(P_\infty)$, where $P_\infty \mid p_\infty$.*

The group-theoretic indices in (3) can be evaluated in some special cases.

([51]) That is to say, the coset belonging to each σ induces in ϵ_i the identity automorphism.

([52]) In this case the p-primary ideals form an infinite cyclic group.

If Γ is abelian and is the direct product of cycles of orders $m_1, \cdots, m_t$, then(53)

$$[FH:TH] = \prod_{i<j} (m_i, m_j). \tag{4}$$

Exactly the same formula holds for an abstract ramified group P, as will be proved by the computations(54) of §22. These results prove

THEOREM 21.2. *Let Γ be an abelian group represented as the direct product of t cyclic groups of respective orders $m_1, m_2, \cdots, m_t$. Then*

$$[FE^*:TE^*] = [\Gamma:\Gamma^r]^r \left[\prod_{i<j} (m_i, m_j)\right]^{r+1} \tag{5}$$

where (m_i, m_j) denotes the greatest common divisor of m_i and m_j.

22. **The number of extensions for a unit group.** This section is devoted to group-theoretic computations leading to

THEOREM 22.1. *If P is an abstract ramified unit group belonging to an abelian group Γ which is the direct product of cyclic groups of orders $m_0, \cdots, m_s$, then the number of group extensions of P by Γ is*

$$[FP:TP] = \prod_{i<j} (m_i, m_j); \qquad i, j = 0, \cdots, s. \tag{1}$$

The ramified unit group P is defined with reference to an automorphism δ of order 2 in Γ (δ is the generator of the decomposition group of a corresponding infinite prime divisor of K). Write the abelian group Γ as a direct product of $s+1$ cyclic groups $\{\alpha_i\}$,

$$\Gamma = \{\alpha_0\} \times \{\alpha_1\} \times \cdots \times \{\alpha_s\}; \qquad \text{order of } \alpha_i = m_i. \tag{2}$$

A simple computation shows that this representation may be so modified that the given δ lies in one of the cycles, as

$$\delta = \alpha_0^r, \qquad m_0 = 2r. \tag{3}$$

In each coset of Γ modulo $\{1, \delta\}$ there is a representative of the form

$$\beta = \alpha_0^{e_0}\alpha_1^{e_1} \cdots \alpha_s^{e_s}; \qquad 0 \leqq e_0 < r, 0 \leqq e_i < m_i, (i \neq 0), \tag{4}$$

with an exponent less than r for α_0.

By definition, the abelian group P is generated by elements P^β subject

(53) Dr. A. H. Clifford has pointed out to us that this product is exactly the order of the multiplicator $\mathfrak{M}$ of Γ in Schur's theory of collineations (see [15, 36]). Moreover, Clifford and MacLane in [12] have proved by more general methods that FH/TH is isomorphic to $\mathfrak{M}$ whenever Γ is solvable.

(54) We omit the analogous computations for (4), since the result is given by [12].

to the relations([55])

$$\prod_{\beta} \mathrm{P}^{\beta} = 1, \qquad \mathrm{P}^{\beta} = \mathrm{P}. \tag{5}$$

Any element A of P can be expressed in the form

$$A = \prod_{\beta} \mathrm{P}^{a(\beta)\beta} = \mathrm{P}^{\sum a(\beta)\beta} = \exp\left(\sum_{\beta} a(\beta)\beta\right), \tag{6}$$

("exp" for convenience!). Since $\mathrm{P}^{\delta} = \mathrm{P}$, it is convenient to define the integer $a(\gamma)$ for all γ in Γ by the convention $a(\beta\alpha_0') = a(\beta\delta) = a(\beta)$. Because of the first relation of (5), the representation (6) is not unique; the exponents $a(\beta)$ may all be changed simultaneously to $a(\beta)+g$, for any integer g. We call g a *change of gauge.*

In any group extension $\mathfrak{E}$ of P by Γ pick a fixed representative $u_i = u(\alpha_i)$ for the cosets α_i of $\mathfrak{E}/\mathrm{P} \simeq \Gamma$. Then use for any element γ of Γ the special representatives generated by the u_i,

$$u(\alpha_0^{e_0}\alpha_1^{e_1}\cdots\alpha_s^{e_s}) = u_0^{e_0}u_1^{e_1}\cdots u_s^{e_s}, \qquad 0 \leqq e_i < m_i. \tag{7}$$

The multiplication table for these representatives is then

$$u_i^{m_i} = C_i, \qquad u_j u_i = D_{ij} u_i u_j \qquad (i < j;\ i, j = 0, \cdots, s); \tag{8}$$

these constants C_i and D_{ij} from P form a *normalized factor set.* The multiplication table for $\mathfrak{E}$ is completed by the commutation rule

$$u_i A = A^{\alpha_i} u_i \qquad \text{(for all } A \text{ in P)}. \tag{9}$$

If the original coset representatives are changed to $v_i = A_i u_i$, for constants $A_0, \cdots, A_s$ in P, the constants C_i and D_{ij} of the factor set are multiplied respectively by the transformation quantities

$$N_i A_i, \qquad A_j^{1-\alpha_i} A_i^{\alpha_j - 1} \qquad (i < j,\ i, j = 0, \cdots, s); \tag{10}$$

here N_i (the "relative norm for α_i") denotes the expression

$$N_i A_i = A^{1+\alpha+\cdots+\alpha^{m-1}}, \qquad (\alpha = \alpha_i,\ m = m_i). \tag{11}$$

The factor set $\{C_i, D_{ij}\}$ must satisfy the associativity conditions([56])

$$C_i^{\alpha_i} = C_i, \qquad C_i^{\alpha_j - 1} = N_i D_{ij}, \qquad D_{ij}^{\alpha_k - 1} D_{jk}^{\alpha_i - 1} D_{ki}^{\alpha_j - 1} = 1 \tag{12}$$

for all i, j, k, with $D_{ji} = D_{ij}^{-1}$ and $D_{ii} = 1$. Our problem is to count the number

([55]) Here and subsequently β denotes an index which runs over the elements (4).

([56]) The fact that these conditions are necessary and sufficient to insure associativity is proved by Zassenhaus [43, p. 97]. His $A_{i,k}$ is our D_{ki}.

of non-equivalent solutions of these equations, where two solutions are called equivalent if their quotient is a transformation set (10). A preliminary is

Step I. Removal of the constants C_i.

LEMMA 1. *For each group extension $\mathfrak{E}$ of $\mathfrak{P}$ by Γ there is a normalized factor set with $C_i=1$, for $i=1, \cdots, s$.*

Proof. If C_i is written in the exponent form of (6), then

$$C_i = \exp\Big(\sum_{\beta} c_i(\beta)\beta\Big), \qquad C_i^{m_i} = \exp\Big(\sum_{\beta} c_i(\alpha_i^{-1}\beta)\beta\Big). \tag{13}$$

According to the associativity condition (12) these two representations can only differ by a change of gauge g_i, so (12) becomes

$$c_i(\alpha_i^{-1}\beta) = c_i(\beta) + g_i \qquad (\text{all } \beta;\ i = 0, \cdots, s). \tag{14}$$

For a fixed i, add these equations over all β. Since $\sum_\beta c_i(\alpha_i^{-1}\beta) = \sum_\beta c_i(\beta)$, each $g_i = 0$. Therefore $c_i(\beta) = c_i(\beta\alpha_i^k)$ for every k. If β' runs over the elements β of (4) which do not involve α_i (i.e., which have $e_i = 0$), the C_i of (13) becomes

$$C_i = \exp\Big[\sum_{\beta'} (\beta' + \beta'\alpha_i + \cdots + \beta'\alpha_i^{m_i-1})c_i(\beta')\Big]; \qquad (i \neq 0). \tag{15}$$

The sum $(1+\alpha+\cdots+\alpha^{m-1})$ is the exponent involved in the norm N_i of (11), so $C_i = N_i(\exp\sum_{\beta'} c_i(\beta'))$. This means that C_i is a transformation quantity N_iA_i for a suitable A_i, so that this transformation will reduce C_i to 1, q.e.d.

For $i=0$ the argument fails. Since β involves α_0^e only up to $e=r-1$, (15) becomes

$$C_0 = \exp\Big[\sum_{\beta'} (\beta' + \beta'\alpha_0 + \cdots + \beta'\alpha_0^{r-1})c_0(\beta')\Big]; \qquad (\beta' \text{ in } \{\alpha_1, \cdots, \alpha_s\}). \tag{16}$$

The corresponding transformation quantity N_0A_0 is (with $\alpha = \alpha_0$)

$$N_0A_0 = A_0^{1+\alpha+\cdots+\alpha^{m-1}} = A_0^{(1+\alpha+\cdots+\alpha^{r-1})(1+\alpha^r)} = A_0^{2(1+\alpha+\cdots+\alpha^{r-1})},$$

for $\alpha^r = \delta$ leaves A_0 fixed. The exponents in N_0A_0 are all even, so C_0 of (16) can be brought to this form only if its exponents $c_0(\beta')$ become even after a change of gauge. This proves

LEMMA 2. *The constant C_0 may be reduced to 1 by a transformation quantity if and only if its exponents $c_0(\beta')$ satisfy $c_0(\beta') \equiv c_0(1) \pmod 2$ for every β' in the group generated by $\alpha_1, \cdots, \alpha_s$.*

Henceforth we consider only factor sets reduced, as in Lemma 1, to

$$C_1 = C_2 = \cdots = C_s = 1, \qquad N_1A_1 = \cdots = N_sA_s = 1. \tag{17}$$

Step II. Construction and counting of the invariants. The essential device for our computation is the reduction of factor sets in P to factor sets of integers. This is done by using the *trace* of each element A of the group P,

$$a = t(A) = t\left\{\exp \sum_{\beta} a(\beta)\beta\right\} \equiv \sum_{\beta} a(\beta) \pmod{n/2}, \tag{18}$$

where $n = m_0 m_1 \cdots m_s$ is the order of the group Γ. The change of gauge $a(\beta) \to a(\beta) + g$ changes the sum over $n/2$ terms β by $(n/2)g$, hence $t(A)$ is uniquely determined modulo $n/2$. Furthermore

$$t(AB) \equiv t(A) + t(B), \qquad t(A^\gamma) \equiv t(A) \pmod{n/2}. \tag{19}$$

Therefore t is an operator-homomorphism of the group P of units on the additive group of integers modulo $n/2$, so t maps the given factor set $\{C_0, D_{ij}\}$ onto a factor set of integers

$$c_0 \equiv t(C_0), \qquad d_{ij} \equiv t(D_{ij}) \qquad (i < j,\ i, j = 0, \cdots, s). \tag{20}$$

The homomorphism t applied to the associativity conditions (12) yields 0 for every term of the form $B^{\alpha-1}$, because $t(B^\alpha) = t(B)$. There remain only the conditions derived from the norms in (12). The norm $N_i B$ of (11) contains m_i exponents, so $t(N_i B) = m_i t(B)$, and the associativity conditions (12) yield the trace conditions

$$m_i d_{ij} \equiv 0 \pmod{n/2}, \qquad m_j d_{ij} \equiv 0 \pmod{n/2}. \tag{21}$$

These must hold for all $i \neq j$, with $d_{ji} = -d_{ij}$.

LEMMA 3. *The number of integers d_{ij} modulo $n/2$ which satisfy a trace condition* (21) *is the g.c.d.* (m_i, m_j).

Case 1: $0 < i < j$. Here (21) states that d_{ij} is a multiple of the integers $n/2m_i$ and $n/2m_j$. The smallest such multiple is the l.c.m. $[n/2m_i, n/2m_j]$; the the number of multiples (mod $n/2$) is

$$(n/2)[n/2m_i, n/2m_j]^{-1} = m_i m_j n [m_j n, m_i n]^{-1} = (m_i, m_j),$$

according to the relation $(m_i, m_j)[m_i, m_j] = m_i m_j$.

Case 2. $i = 0$, $n/m_0 \equiv 0 \pmod 2$. Here $n/2m_0$ is an integer, so the argument proceeds exactly as in Case 1.

Case 3. $i = 0$, $n/m_0 \not\equiv 0 \pmod 2$. The congruence $m_0 d_{0j} = 2r d_{0j} \equiv 0$ of (21) becomes here $2d_{0j} \equiv 0 \pmod{n/m_0}$. The latter modulus is odd by assumption, so $d_{0j} \equiv 0 \pmod{n/m_0}$. Computations now give (m_j, r) solutions d_{0j}. But the hypothesis $n/m_0 \not\equiv 0 \pmod 2$, where $n/m_0 = m_1 \cdots m_j$, means that each m_j must be odd. Therefore $(m_j, r) = (m_j, 2r) = (m_j, m_0)$ is the number of solutions.

Because this lemma gives the same total count $\prod_{i<j}(m_i, m_j)$ as that stated in our theorem, it will suffice to prove that each set of solutions d_{ij} of the trace

conditions (21) is the trace belonging to one and only one group extension. Because a transformation quantity D_{ij} of (10) has trace zero, each group extension $\mathfrak{E}$ determines the trace d_{ij} uniquely. It remains only to prove that each set of solutions d_{ij} is the trace of some factor set and that a factor set with traces $d_{ij}=0$ is a transformation quantity (and so belongs to the unit extension).

Step III. The realization of invariants by factor sets. To given d_{ij} we wish to construct a factor set

$$C_0 = \exp\left[\sum_{\beta} c_0(\beta)\beta\right], \qquad D_{ij} = \exp\left[\sum_{\beta} d_{ij}(\beta)\beta\right]. \tag{22}$$

The integral exponents $c_0(\beta)=c(\beta)$ and $d_{ij}(\beta)$ must satisfy the associativity conditions (12). The second condition of (12), after substitution of the expressions (22), takes on the forms (the first for $i=0$, the second for $i\neq 0$)

$$g_{0j} + c_0(\beta\alpha_j^{-1}) - c_0(\beta) = 2\sum_{k=0}^{r-1} d_{0j}(\beta\alpha_0^k), \qquad (j \neq 0, \text{ all } \beta), \tag{23}$$

$$g_{ij} = \sum_{k=0}^{m_i-1} d_{ij}(\beta\alpha_i^k), \qquad (\text{all } j \neq i,\ i \neq 0), \tag{24}$$

where g_{ij} represents a change of gauge. These gauges may be determined in terms of the traces. In (24) let β' run through all elements (4) not involving α_i. By adding these $n/2m_i$ equations, one gets

$$ng_{ij}/2m_i = \sum_{\beta'}\sum_{k=0}^{m_i-1} d_{ij}(\beta'\alpha_i^k) = \sum_{\beta} d_{ij}(\beta) = d_{ij}.$$

Hence $g_{ij}=2m_id_{ij}/n$. A similar addition of (23) gives g_{0j}. For a fixed choice of gauge for C_0 and D_{ij} in (22) we have

$$g_{ij} = 2m_id_{ij}/n, \qquad (\text{all } i \neq j). \tag{25}$$

The trace conditions (21) assert that these constants g_{ij} are integers.

Since $t(AB)=t(A)+t(B)$, it will be possible to realize any allowable set of invariants d_{ij} if these invariants can be realized one at a time. The cases $i=0$ and $0<i<j$ behave differently, as we now show by treating the typical cases d_{12} and d_{01}.

LEMMA 4. *If d_{12} is a solution of the trace conditions* (21), *with $i=1$, $j=2$, there exists a factor set $\{C_0, D_{ij}\}$ with $C_0=D_{ij}=1$ for $(i, j)\neq(1, 2)$ and with $t(D_{12})=d_{12}$.*

Proof. The essential associativity condition is (24) with $i=1$, $j=2$ and $i=2, j=1$; the trace conditions insure that g_{12} and g_{21} are integers. If we regard the exponents $d_{12}(\sigma\alpha_1^h\alpha_2^k) = -d_{21}(\sigma\alpha_1^h\alpha_2^k)$ for fixed σ as the terms of an m_1 by m_2

matrix, condition (24) requires that each column ($h=0, \cdots, m_1-1$) add up to g_{12}, and that each row ($k=0, \cdots, m_2-1$) add up to $-g_{21}$. These two requirements are consistent because the sum of all terms by rows is $-m_1g_{21}$, by columns is m_2g_{12}, and these two sums are equal by (25), with $d_{12}=-d_{21}$. The requirements may be met by making the matrix $\|d_{12}(\sigma\alpha_1^h\alpha_2^k)\|$ zero except in the first row and column, as

$$d_{12}(\sigma\overset{h}{\alpha_1}\overset{k}{\alpha_2}) = 0,\ d_{12}(\sigma\overset{h}{\alpha_1}) = -g_{21},\ d_{12}(\sigma\overset{k}{\alpha_2}) = g_{12}\ (h, k \neq 0),$$

$$d_{12}(\sigma) = g_{12} + g_{21} + m_1g_{21} = g_{12} + g_{21} - m_2g_{12}.$$

For any $\sigma=\alpha_0^{e_0}\alpha_3^{e_3}\cdots\alpha_s^{e_s}$ in the group generated by the remaining α's we take these equations as defining the quantity D_{12}. They satisfy the associativity conditions (24), as may be seen by substitution. Furthermore this definition is independent of the choice of σ, so the resulting D_{12} is invariant under any $\alpha\neq\alpha_1$ and α_2. Therefore $D_{12}^{a-1}=1$. Every other D_{ij} is to be 1, so the last associativity condition connecting any three D's as in (12) is satisfied, and the lemma holds.

LEMMA 5. *If d_{01} is a solution of the trace conditions* (21) *with $i=0$ and $j=1$, there exists a factor set $\{C_0, D_{ij}\}$ with $D_{ij}=1$ for $(i, j)\neq(0, 1)$ and with $t(D_{01})=d_{01}$.*

Proof. Much as in the proof of Lemma 3, a separate treatment is necessary in the case when $m_1\equiv 0 \pmod 2$ (i.e., when α_1 has an even order).

Case 1. $m_1\not\equiv 0 \pmod 2$. If we set $C_0=1$, the associativity conditions ((24) with $j=0$, $i=1$ and (23) with $j=1$) become

$$\sum_{h=0}^{m_1-1} d_{01}(\sigma\overset{h}{\alpha_0}\overset{k}{\alpha_1}) = -g_{10}, \qquad \sum_{h=0}^{r-1} d_{01}(\sigma\overset{h}{\alpha_0}\overset{k}{\alpha_1}) = g_{01}/2.$$

By (25) the integers g_{10} and g_{01} satisfy $m_1g_{01}=-m_0g_{10}$. Since $m_0=2r$ is even and m_1 is odd, g_{01} must be even and $g_{01}/2$ is an integer. These equations again specify the row and column sums of a matrix, and may be solved exactly as in the previous lemma.

Case 2. $m_1\equiv 0 \pmod 2$. We choose $d_{10}=-d_{01}$. By the trace condition on d_{01}, the constants g_{01} and g_{10} of (25) are integers, and by (25) they satisfy $m_1g_{01}=-m_0g_{10}$. We no longer choose $C_0=1$, but set instead

$$C_0(\sigma\overset{h}{\alpha_0}\overset{k}{\alpha_1}) = kg_{01}, \tag{26}$$

$$(\sigma = \alpha_2^{e_2}\cdots\alpha_s^{e_s};\ h = 0, \cdots, r-1;\ k = 0, \cdots, m_1-1).$$

For $\beta=\sigma\alpha_0^h\alpha_1^k$ the quantity $c_0(\beta\alpha_1^{-1})-c_0(\beta)$ of (23) is, according to this agreement, $(k-1)g_{01}-kg_{01}=-g_{01}$, provided $k\neq 0$. If $k=0$, it is $(m_1-1)g_{01}$. Therefore (23) and (24) give the associativity conditions

$$(27)\qquad \sum_{h=0}^{r-1} d_{01}(\sigma\alpha_0^h) = m_1 g_{01}/2, \qquad \sum_{h=0}^{r-1} d_{01}(\sigma\alpha_0^h\alpha_1^k) = 0, \qquad (k \not\equiv 0),$$

$$(28)\qquad \sum_{h=0}^{m_1-1} d_{01}(\sigma\alpha_0^h\alpha_1^k) = -g_{10}, \qquad (\sigma = \alpha_2^{e_2}\cdots\alpha_s^{e_s}).$$

These requirements once more specify consistent row and column sums for a matrix $\|d_{01}(\sigma\alpha_0^h\alpha_1^k)\|$. A solution is given by

$$(29)\qquad d_{01}(\sigma\alpha_0^h\alpha_1^k) = 0, \quad (k \not\equiv 0); \qquad d_{01}(\sigma\alpha_0^h) = -g_{10},$$

where σ again is any element in the subgroup generated by $\alpha_2, \cdots, \alpha_s$. These formulas provide a construction of a D_{01} from the given integer d_{01}. Because the formulas (29) do not depend on σ, $D_{01}^{\alpha} = D_{01}$ for any α not α_0 or α_1. For this reason all the associativity conditions of (12) are satisfied if the remaining D_{ij} are all 1. This proves the lemma. Combination of these results proves

LEMMA 6. *Any set of integral solutions d_{ij} of the trace conditions* (21) *is the trace of some factor set* $\{C_0, D_{ij}\}$.

In Case 2 of the proof of Lemma 5 the use of the constant $C_0 \not\equiv 1$ is essential (at least whenever $g_{01} \not\equiv 0 \pmod 2$). One may show by examples that this is the case; the result will subsequently throw some light on the relation of the ramified group P to Schur's multiplicator.

LEMMA 7. *For some abelian groups* Γ *there is a factor set* $\{C_0, D_{ij}\}$ *with* $t(C_0) \not\equiv 0 \pmod{n/2}$.

Proof. The expression C_0 constructed in (26) has

$$t(C_0) \equiv (n/2)(m_1 + 1)g_{01}/2.$$

This will satisfy $t(C_0) \equiv 0 \pmod{n/2}$ only if m_1 is odd or g_{01} is even. Sometimes neither will be the case, as for instance when $\Gamma = \{\alpha_0\} \times \{\alpha_1\}$ with α_0 of order 2, α_1 of order 6. Then $m_1 = 6 \equiv 0 \pmod 2$, $n = 12$, and the constant $d_{01} = 3$ satisfies the trace condition. It gives $g_{01} = 1$ by (25). Hence $t(C_0) = 6(7/2) = 21 \not\equiv 0 \pmod 6$. Since $t(C_0)$ cannot be altered, mod 2, by the insertion of any transformation quantities (10), the factor set so constructed cannot be equivalent to any factor set with $C_0 = 1$.

Step IV. The reduction of factor sets with zero traces.

LEMMA 8. *If a factor set* $\{C_0, D_{ij}\}$ *has traces* $t(D_{ij}) = d_{ij} \equiv 0 \pmod{n/2}$, *then C_0 may be reduced to* 1 *by a suitable transformation set*

Proof. By suitable choice of gauge in (22), all d_{ij} are zero identically. Then $g_{0j} = 0$ by (25), so the associativity condition (23) makes $c_0(\beta\alpha_j^{-1}) - c_0(\beta) \equiv 0 \pmod 2$. By applying this repeatedly for all $j = 1, \cdots, s$, one proves

$c_0(\beta') \equiv c_0(1) \pmod 2$ for every β' in the group generated by $\alpha_1, \cdots, \alpha_s$. According to Lemma 2, C_0 may then be reduced to 1.

After this reduction the associativity conditions (12) satisfied by any one constant D_{ij} have the form

$$N_i D_{ij} = 1, \qquad N_j D_{ij} = 1, \qquad t(D_{ij}) \equiv 0 \pmod{n/2}. \tag{30}$$

The essential step is the demonstration that each D which satisfies conditions of this form may be removed by an appropriately chosen transformation quantity. This is essentially accomplished by a lemma related closely to a lemma of Clifford's [12].

LEMMA 9. *If an element D in the unit group P satisfies conditions $t(D) \equiv 0 \pmod{n/2}$ and $N_i D = 1$, there exists in P an element A such that $D = A^{1-\alpha}$, where $\alpha = \alpha_i$. The quantity A may be so constructed that, for every $j \neq i$, $N_j D = 1$ implies $N_j A = 1$ and $D = D^{\alpha^j}$ implies $A = A^{\alpha^j}$.*

Proof. Write $D = \exp\left(\sum_\beta d(\beta)\beta\right)$. The hypothesis $t(D) \equiv 0$ means that $d = \sum_\beta d(\beta) \equiv 0 \pmod{n/2}$. By a suitable choice of gauge for D we may actually get $d = 0$. The hypothesis $N_i D = 1$ may now be expressed in terms of the exponents $d(\beta)$ exactly as in (23) and (24). Because the trace d is zero, the gauges g_{ij}, computed as in (25), are also zero, so the condition $N_i D = 1$ becomes

$$\sum_h d(\beta\alpha_i^h) = 0 \qquad (h = 0, \cdots, m_i - 1). \tag{31}$$

We seek an $A = \exp\left\{\sum_\beta a(\beta)\beta\right\}$ to satisfy the equation $A^{1-\alpha} = D$, where $\alpha = \alpha_i$. In terms of the exponents $a(\beta)$ this equation is

$$a(\beta) - a(\beta\alpha^{-1}) = d(\beta) + g \qquad (g \text{ a constant for all } \beta). \tag{32}$$

To solve this explicitly, write $\beta = \tau\alpha^e$ in such wise that τ involves only the generators $\alpha_j \neq \alpha$, and try

$$a(\tau\alpha^e) = \sum_{k=0}^{e} d(\tau\alpha^k), \qquad (e = 0, 1, \cdots, m_i - 1, \text{ if } i \neq 0). \tag{33}$$

If $e = m_i$, then $\tau\alpha^e = \tau$ and this equation is still valid in virtue of the hypothesis (31). Therefore (33) holds for every integer $e \geqq 0$. By substitution we then find that this proposed A does satisfy the required condition (32), with a gauge $g = 0$.

The case $\alpha = \alpha_0$ may be treated by a minor modification of this argument. For a solution A we need define the exponent $a(\tau\alpha^e)$ only in the range $e = 0, \cdots, r-1$, where $r = m_0/2$. Again we adopt (33) as the definition in this range. Since the original exponents d satisfy $d(\beta\alpha^r) = d(\beta)$, the hypothesis (31) may now be written $2\sum_h d(\beta\alpha_0^h) = 0$, where h runs from 0 to $r-1$. There-

fore (33) again holds for all integers e outside the original range of definition, and we have constructed A to satisfy $a(\beta\alpha^r)=a(\beta)$, as desired.

The construction (33) gives an element A with the added properties asserted in the lemma. Because $t(D)=0$, an added hypothesis $N_jD=1$ would again mean

$$\sum_h d(\beta\alpha_j^h) = 0, \qquad (h = 0, 1, \cdots, m_j - 1). \tag{34}$$

On the other hand, N_jA is by definition

$$N_jA = \exp\left\{\sum_\beta a(\beta)\beta\left(\sum_h \alpha_j^h\right)\right\} = \exp\left\{\sum_\beta\left(\sum_{h'} a(\beta\alpha_j^{h'})\beta\right)\right\}.$$

Set $\beta=\tau\alpha^s$, and compute each exponent, by (33), as

$$\sum_h a(\tau\alpha^s\alpha_j^h) = \sum_{h=0}^{m_j-1}\sum_{k=0}^{d} d(\tau\alpha^k\alpha_j^h) = \sum_k\sum_h d(\tau\alpha^k\alpha_j^h).$$

This sum is 0, by the hypothesis (34); hence $N_jA=1$.

Suppose next that $D=D^{\alpha_j}$. Exactly as in the argument in (14), this means that the exponents satisfy $d(\beta)=d(\beta\alpha_j)$. With the definition (33) one then has also $a(\beta)=a(\beta\alpha_j)$, hence $A=A^{\alpha_j}$, q.e.d.

By this lemma, any individual D_{ij} may be reduced. For example $N_0D_{01}=N_1D_{01}=1$ by the associativity conditions (30), so there must exist an A_0 with $A_0^{1-\alpha_1}=D_{01}$ and $N_0A_0=1$. According to the formulas (10) for transformation quantities, a change $v_0=A_0^{-1}u_0$ in the coset representation will then replace D_{01} by 1 without disturbing the normalization $C_0=1$ already achieved in Lemma 8. This process continues as in

LEMMA 10. *If $1\leqq k\leqq s$, every factor set $\{C_i, D_{ij}\}$ with traces $t(D_{ij})\equiv 0$ is similar to a factor set with $C_i=1$ for all i and $D_{ij}=1$ for all i and j which satisfy $0\leqq i<j\leqq k$.*

Proof. The above reduction of D_{01} takes care of the case $k=1$. Assume by induction that the lemma is true for $k-1$, and by a second induction that $D_{0k}=\cdots=D_{j-1\,k}=1$ (the case $j=0$ will be included in the induction argument). For the next constant D_{ik} use at long last the third associativity condition of (12), which enforces a relation between three constants D. For each $i<j$ the constants D_{ij} and $D_{ki}=D_{ik}^{-1}$ have already been made equal to 1, so this third condition becomes $D_{jk}^{\alpha-1}=1$ for $\alpha=\alpha_0, \cdots, \alpha_{j-1}$. The other conditions on D_{jk} are $N_jD_{jk}=N_kD_{jk}=1$. By Lemma 9 there exists an A_k with

$$D_{jk} = A_k^{1-\alpha_j}, \qquad N_kA_k = 1, \qquad A_k^{\alpha_0} = \cdots = A_k^{\alpha_{j-1}} = A.$$

Make the corresponding change $v_k=A_k^{-1}u_k$ in the coset representatives (7). According to the list of transformation quantities (10), this change reduces

D_{jk} to 1 and modifies each of C_k and D_{ik} for $i=0, \cdots, j-1$ by 1. In other words, the new reduction $D_{jk}=1$ is accomplished without disturbance of the results of previous reductions. This completes the induction, and proves

LEMMA 11. *Every factor set with traces zero is similar to* 1.

With this lemma the proof of Theorem 1 of this section is complete. Our argument has actually shown that the group of classes of factor sets is isomorphic to the additive group of traces d_{ij}. Each trace by itself gives a cyclic group of order (m_i, m_j), hence

THEOREM 22.2. *The group* $F\mathrm{P}/T\mathrm{P}$ *of group extensions of* P *by an abelian group* Γ *is the direct product of cyclic groups of orders* (m_i, m_j), *for all* $0 \leqq i < j \leqq s$.

For an unramified unit group H exactly the same two theorems are true; they have been established by more general arguments than ours in Clifford [12]. It is proved there that FH/TH is isomorphic to the multiplicator $\mathfrak{M}$ of Γ in an algebraically closed field of characteristic ∞. This multiplicator can be represented by factor sets of nth roots of unity or, equally well, of integers modulo n, but the resulting factor sets of integers can always be normalized so that the constants for $c_i = u_i^{m_i}$ are all 0. In the unramified case, the trace maps the group FH/TH isomorphically on this representation of the multiplicator $\mathfrak{M}$, as is proved by Clifford. In the ramified case the group $F\mathrm{P}/T\mathrm{P}$ is still isomorphic to the multiplicator, in virtue of Theorem 22.2; however, the trace $t(A)$ no longer provides an isomorphic map of $F\mathrm{P}/T\mathrm{P}$ on $\mathfrak{M}$. This curious anomaly is a consequence of Lemma 7 above, which asserts that for the ramified case the constant C_0 cannot always be reduced to 1, although it can be so reduced in the case of the multiplicator.

23. **Crossed characters for abstract unit groups.** The index $J(E^*)$ for the Herbrand unit group E^* has as numerator $[UE^*:E^{*1-\sigma}]$, the number of classes of crossed characters (see §7) of Γ in E^*. Since E^* is a direct product (see the end of §17), this index can be written as a product

$$(1) \qquad [UE^*:E^{*1-\sigma}] = \prod_{i=1}^{r} [U\epsilon_i:\epsilon_i^{1-\sigma}] \prod_{i=1}^{r+1} [UH_i:H_i^{1-\sigma}].$$

The first r factors in (1) involve the number of classes of crossed characters of Γ in an infinite cyclic group ϵ_i generated by one of the r chosen independent units ϵ_i of $\epsilon \cap E^*$. Since the elements of ϵ are invariant under the operators of Γ, it results that these crossed characters are all ordinary characters. But Γ is a finite group, hence it has no non-trivial ordinary characters in an infinite cyclic group. The first product in (1) is therefore 1.

The second product in (1) involves both ramified and unramified unit groups H_i. Consider for the moment a ramified group P, associated with an automorphism δ of order 2, which generates the corresponding decomposition

group. As in definition (10) of §17, the group P then has $n/2$ generators P^{σ_i}, where σ_i are representatives of the left cosets of Γ modulo $\{1, \delta\}$. An element A of the group P can be represented (but not uniquely!) in the form

$$A = P^{a_1\sigma_1}P^{a_2\sigma_2}\cdots P^{a_m\sigma_m}, \qquad m = n/2. \tag{2}$$

We again associate with A the "trace" as an essential invariant([57]),

$$a = t(A) \equiv a_1 + a_2 + \cdots + a_m \pmod{m}. \tag{3}$$

Since the representation (2) of A can be altered only by using the relation $1 = P^{\sigma_1}\cdots P^{\sigma_m}$, the function $t(A)$ is uniquely determined by A, modulo m. Furthermore

$$\begin{aligned} t(AB) &\equiv t(A) + t(B) \pmod{m}, \\ t(A^\tau) &\equiv t(A) \pmod{m}, \qquad \tau \text{ in } \Gamma. \end{aligned} \tag{4}$$

Given a crossed character $U(\sigma)$ of Γ in this group P, we define

$$\chi(\sigma) = t[U(\sigma)], \qquad \text{all } \sigma \text{ in } \Gamma. \tag{5}$$

The definition of a crossed character (§7) implies at once that this function is an (ordinary) character of Γ in the group of integers modulo m, with

$$\chi(\sigma\tau) \equiv \chi(\sigma) + \chi(\tau) \pmod{m}. \tag{6}$$

We say that a crossed character $U\mathrm{P}$ lies in the principal genus $G\mathrm{P}$ if and only if the corresponding function χ is the identity, so that the principal genus consists of all functions $G(\sigma)$ on Γ to P with the properties

(*23.1) $G(\sigma\tau) = G(\sigma)[G(\tau)]^\sigma$, $t(G_\sigma) \equiv 0 \pmod{m}$.

By using the homomorphism which carries each crossed character U into the character χ of (5), we shall prove

THEOREM 23.1. *For a ramified abstract unit group* P *whose decomposition group is generated by* δ $(\delta^2 = 1)$ *the number of classes of crossed characters is given by*

$$[U\mathrm{P}:\mathrm{P}^{1-\sigma}] = [G\mathrm{P}:\mathrm{P}^{1-\sigma}][\Gamma:\{\Gamma', \delta\}], \tag{7}$$

where the first index measures the number of classes of crossed characters in the principal genus, while the second factor is the index in Γ *of the subgroup generated by* δ *and the commutator group* Γ'.

Proof. The index $[\Gamma:\{\Gamma', \delta\}]$ is simply the number of (ordinary) characters of the abelian factor group $\Gamma/\{\Gamma', \delta\}$. The homomorphism principle of §9 applied to the map $U \to \chi$ will therefore yield the conclusion (7) directly,

([57]) The character group of P in the group of real numbers modulo 1 is a cyclic group generated by $\chi(A) = t(A)/m$. This is the reason for the importance of t.

once we prove that the characters χ obtained by this homomorphism are indeed the characters of $\Gamma/\{\Gamma', \delta\}$. To achieve this goal we prove two lemmas, from which the desired description of the χ will follow immediately.

LEMMA 1. *For any σ in $\{\Gamma', \delta\}$, $\chi(\sigma) \equiv 0 \pmod{m}$.*

Proof. This lemma asserts that χ can be considered as a character defined for the cosets of the factor group $\Gamma/\{\Gamma', \delta\}$. First observe that (6) shows that $\chi(\sigma) \equiv 0 \pmod{m}$ whenever σ is a commutator in Γ'. It remains only to prove $\chi(\delta) \equiv 0 \pmod{m}$. By the definition (5), $\chi(\delta) \equiv t[U(\delta)]$. As in (2), we use for $U(\delta)$ the notation

$$U(\delta) = \prod_{i=1}^{m} \mathrm{P}^{d_i \sigma_i}, \qquad \chi(\delta) \equiv \sum_{i=1}^{m} d_i \pmod{m}. \tag{8}$$

For a crossed character, $U_\sigma U_\tau^\sigma = U_{\sigma\tau}$, by definition. With $\sigma = \tau = 1$ and $\sigma = \tau = \delta$ this gives the conclusions

$$U(1) = 1, \qquad [U(\delta)]^{1+\delta} = U(\delta^2) = 1. \tag{9}$$

We proceed to express the second of these conditions in terms of the exponents d_i entering into the value $\chi(\delta)$ of (8). For

$$[U(\delta)]^\delta = \prod_{i=1}^{m} (\mathrm{P}^{d_i \sigma_i})^\delta = \prod_{i=1}^{m} \mathrm{P}^{d_i \delta \sigma_i} = \prod_{i=1}^{m} \mathrm{P}^{d_i \sigma_j}$$

where σ_j $(j = j_i)$ is the representative belonging to the left coset $\sigma_j\{1, \delta\}$ in which the product $\delta\sigma_i$ lies. The correspondence $i \leftrightarrow j_i = j$ is an involution, so $\sum d_i \sigma_j = \sum d_j \sigma_i$ and

$$[U(\delta)]^{1+\delta} = \prod_{i=1}^{m} \mathrm{P}^{d_i \sigma_i} \prod_{i=1}^{m} \mathrm{P}^{d_i \sigma_j} = \prod_{i=1}^{m} \mathrm{P}^{s_i \sigma_i}$$

where $s_i = d_i + d_j$. This product is 1 by (9), and this must be a consequence of the basic relation $\prod_{i=1}^{m} \mathrm{P}^{\sigma_i} = 1$ for our abstract ramified unit group P. There is therefore an integer g such that

$$d_i + d_j = g, \qquad i = 1, 2, \cdots, m, \; j = j_i. \tag{10}$$

If $i = 1$ belongs to the coset of 1, j_i is also 1. Consequently $g = d_1 + d_1 = 2d_1$ is even. If i runs over all integers from 1 to m, so does j_i. Adding (10) over all i, we get

$$2\sum_{i=1}^{m} d_i = gm, \qquad \sum_{i=1}^{m} d_i = (g/2)m \equiv 0 \pmod{m}.$$

According to the formula (8) for $\chi(\delta)$, this yields $\chi(\delta) \equiv 0 \pmod{m}$, as required for the lemma.

LEMMA 2. *Any character* χ *of* $\Gamma/\{\Gamma', \delta\}$ *can be realized as a trace* (5) *of some crossed character* $U\mathrm{P}$.

Proof. We observe first that δ has order 2, whence $\{\Gamma', \delta\}$ has even order and $\Gamma/\{\Gamma', \delta\}$ has as order a divisor of $m=n/2$. This means that any one of the $[\Gamma:\{\Gamma', \delta\}]$ characters of the abelian quotient group $\Gamma/\{\Gamma', \delta\}$ can in fact be expressed as a homomorphic mapping $\sigma\rightarrow\chi(\sigma)$ of Γ on the integers modulo m. As is well known, the representation of the abelian group $\Gamma/\{\Gamma', \delta\}$ as a direct product of cyclic groups makes it possible to write any character of $\Gamma/\{\Gamma', \delta\}$ as a product of characters, each of which is defined on a cyclic quotient group of Γ. To consider such a sample quotient group, let $\Omega\supset\{\Gamma', \delta\}$ be a subgroup of Γ such that Γ/Ω is cyclic of order h, and let α be a representative of some coset of Γ/Ω which generates this cyclic group. Then any σ in Γ can be written as

$$\sigma = \alpha^i\omega; \qquad i = 0, 1, \cdots, h-1; \ \omega \text{ in } \Omega, \ \alpha^h \text{ in } \Omega. \tag{11}$$

To realize all possible characters χ of Γ by crossed characters it will thus suffice to realize any character of the form

$$\begin{aligned} &\chi(\alpha) \equiv m/h \pmod m, \\ &\chi(\omega) \equiv 0 \pmod m, \qquad \text{all } \omega \text{ in } \Omega. \end{aligned} \tag{12}$$

Let Ω be decomposed modulo $\{1, \delta\}$ into m/h left cosets, $\Omega=\sum_j\omega_j\{1, \delta\}$, with representatives ω_j. Define next an element C in P by the equations

$$C = \prod_j \mathrm{P}^{\omega_j}, \qquad j = 1, 2, \cdots, m/h. \tag{13}$$

This element C has the properties

$$\begin{aligned} &t(C) \equiv m/h \pmod m, \\ &C^\omega = C, \qquad \text{for } \omega \text{ in } \Omega. \end{aligned} \tag{14}$$

In terms of this element C we propose to define for every group element σ, given as in (11), a value

$$U(\alpha^i\omega) = C^{1+\alpha+\cdots+\alpha^i}, \qquad \sigma = \alpha^i\omega, \tag{15}$$

of a crossed character $U(\sigma)$. This definition is given only for $i=0, \cdots, h-1$, but the formula holds good for all values of i, because α^h lies in Ω and because the definition of C implies that $C^{1+\alpha+\cdots+\alpha^{m-1}}=1$. With the formula (15), so generalized, a straightforward computation then shows that the function $U(\sigma)$ is a crossed character. By (14) and (15)

$$\begin{aligned} &\chi(\alpha) = t[U(\alpha)] = t(C) \equiv m/h \pmod m, \\ &\chi(\omega) = t(C^\omega) \equiv 0 \pmod m. \end{aligned}$$

Therefore this crossed character does realize the sample character χ given in (12). This completes the proof of Lemma 2 and with it, of Theorem 23.1.

It remains to consider the unramified abstract unit groups H. In such a group an arbitrary element $A = \prod_\sigma H^{a(\sigma)\sigma}$ again has a trace, given by

$$t(A) \equiv \sum_\sigma a(\sigma) \pmod{n}. \tag{16}$$

The principal genus consists of those functions G on Γ to H which satisfy

(*23.2) $G(\sigma\tau) = G(\sigma)[G(\tau)]^\sigma$, $t[G(\sigma)] \equiv 0 \pmod{n}$.

The subsequent argument proceeds as in the ramified case, with appropriate simplifications. The result is

THEOREM 23.2. *For an unramified abstract unit group H the number of classes of crossed characters of Γ in H is*

$$[UH:H^{1-\sigma}] = [GH:H^{1-\sigma}][\Gamma:\Gamma'], \tag{17}$$

where GH is the principal genus, Γ' the commutator group.

In many cases this approach will actually yield explicit answers. For a solvable group Γ, A. H. Clifford has shown [12] that the principal genus index in the unramified case is 1,

$$[GH:H^{1-\sigma}] = 1. \tag{18}$$

The authors have extended Clifford's proof to the ramified case, for Γ abelian.

These results can be combined to give the number of crossed characters for the whole Herbrand unit group E^*. For the ramified cases this will involve the index $[\Gamma:\{\Gamma', \delta\}]$ which will have the value $[\Gamma:\Gamma']$ or $[\Gamma:\Gamma']2^{-1}$ according as the automorphism δ of order 2 does or does not lie in the commutator group Γ'. Now δ was introduced as a generator of the decomposition group of a factor P_∞ of one of the ramified infinite prime divisors p_∞ of K/k. The field of elements left invariant by δ is then the largest subfield of K in which p_∞ is unramified. Consequently the statement that δ lies in Γ' is equivalent to the statement that p_∞ is unramified in the field K' corresponding to Γ'. Only in this case is the index $[\Gamma:\{\Gamma', \delta\}]$ equal to $[\Gamma:\Gamma']$. All told, the $r_1 - \rho_1$ ramified unit groups P_i contribute to the crossed character formula (1) a term $[\Gamma:\Gamma']^{r_1-\rho_1} 2^{-\rho'}$, where

(*23.3) $\rho' =$ the number of infinite prime divisors of k which are ramified in K'.

The combination of Theorems 23.1 and 23.2 yields

THEOREM 23.3. *The number of classes of crossed characters for any Herbrand unit group E^* is*

$$(19)\qquad [UE^*:E^{*1-\sigma}] = [\Gamma:\Gamma']^{r+1}[GH:H^{1-\sigma}]^{\rho_1+r_2-\rho'}\prod_{p\infty}[GP:P^{1-\sigma}],$$

where H is an abstract unramified unit group, and where the last product is taken over all the unit groups $P=P_\infty$ belonging to the ramified prime divisors p_∞ of K/k, while GH, GP refer to the principal genus defined in (*23.1), (*23.2).

Chapter IV. The final formula

24. The composite result. The unit index $J(E)$ has now been evaluated in terms of the correction factor $\phi = J(E)/J(E^*)$ and in the explicit computations of Theorems 23.3 and 22.1. If these results are put in the formula (15) of §16, we obtain

Theorem 24.1. *For any normal extension K/k and any module M which involves all prime divisors p of k ramified in K/k, the least common multiple $J=J(\Gamma)$ of the orders of the elements of the Galois group Γ can be expressed as*

$$(1)\qquad J = n^*n^{-1}\omega(A', E)\omega(\mathfrak{A}', (A'))^{-1}2^{\rho}J(A\mathfrak{f})^{-1}J(E),$$

where $J(\mathfrak{A}^{\mathfrak{f}}) = [U\mathfrak{A}^{\mathfrak{f}}:\mathfrak{A}^{\mathfrak{f}1-\sigma}][F\mathfrak{A}^{\mathfrak{f}}:T\mathfrak{A}^{\mathfrak{f}}]^{-1}$ depends only on the structure of the class group $\mathfrak{A}^{\mathfrak{f}}$, while $J(E)$ depends only on the structure of the group of units in K, by

$$(2)\qquad J(E) = \phi[\Gamma:\Gamma']2^{-\rho'}\left\{[GH:H^{1-\sigma}][FH:TH]^{-1}\right\}^{\rho_1+r_2} \cdot \prod_{i=1}^{\rho}\left\{[GP_i:P_i^{1-\sigma}][FP_i:TP_i]^{-1}\right\}.$$

Alternately, $J(E)$ may be expressed by the formula of Theorem 20.1. *All the indices appearing in these formulas are finite.*

We pause to identify the various invariants which appear in these formulas. In J, n^* is an integer which is a divisor of n and a multiple of $J(\Gamma)$, as given explicitly by formula (*13.4);

Γ' = the commutator group of Γ;

ρ' = the number of infinite prime divisors in k ramified in the commutator subfield K' of K;

$\mathfrak{A}^{\mathfrak{f}}$ = the groups of ideal classes in K;

$A'(\mathfrak{A}')$ = the groups of numbers (ideals) in K which are relatively prime to M.

The class group enters in the numerator only in terms of $[F\mathfrak{A}^{\mathfrak{f}}:T\mathfrak{A}^{\mathfrak{f}}]$, the number of extensions of $\mathfrak{A}^{\mathfrak{f}}$ by Γ, in the denominator only in terms of $[U\mathfrak{A}^{\mathfrak{f}}:\mathfrak{A}^{\mathfrak{f}1-\sigma}]$, the number of classes of crossed characters (see §7) of $\mathfrak{A}^{\mathfrak{f}}$. The groups A', $\mathfrak{A}'$

occur only in connection with the deficiency invariants which were defined in §15.

In the unit index (2) the correction factor $J(E)J(E^*)^{-1}=\phi$ is the quantity considered in §19. It need not be 1, as it would be in the cyclic case. It depends on the position of the Herbrand unit group E^* within the whole group E of units in K.

ρ = the number of infinite prime divisors p of k ramified in K;

ρ_1+r_2 = the number of infinite prime divisors p of k which are unramified in K;

H = the abstract unramified unit group for Γ;

P_i = the abstract ramified unit groups belonging to δ_i: the generator of a decomposition group for the ith ramified infinite prime divisor of k.

It should be noted that the group H, as described in §17, depends only on the Galois group Γ of K/k and not on the further structure of K/k, while the various groups P_i depend only on Γ and the position in Γ of the automorphisms δ_i. Each δ_i may be described as an automorphism in Γ of order 2, such that the corresponding ramified infinite prime divisor $p_{\infty,i}$ of k has a factor P_i which is unramified in the subfield of K left invariant by δ_i.

The groups H enter into the formula (2) in two fashions:

(a) through the number $[FH:TH]$ of extensions of Γ by H, and

(b) through the number $[GH:H^{1-\sigma}]$ of classes of crossed characters in the "principal genus" of crossed characters, defined by (*23.2).

These quantities are explicitly computed in the case of abelian groups, as stated in §§22 and 23. Furthermore $\Gamma=\Gamma'$ in these cases.

THEOREM 24.2. *If the Galois group Γ of K/k is an abelian group, the direct product of t cyclic groups of the respective orders m_i, then the least common multiple $J=[m_1, \cdots, m_t]$ is*

$$J = \phi n^*\{\omega(A', E)\omega(\mathfrak{A}', (A'))^{-1}\}J(\mathfrak{A}')^{-1}\left\{\prod_{i<j}(m_i, m_j)\right\}^{-r-1} \tag{3}$$

where $r+1$ is the number of infinite prime divisors of k.

The quantities which appear in our formulas (1) and (2) are all invariants of the extension K/k. This follows at once from the definition of the quantities, except for the quotient $\phi=J(E)J(E^*)^{-1}$ and the deficiency invariants ω, which apparently depend on the module M. That the deficiency invariants ω are in fact independent of M, if M has the properties of Theorem 24.1, was shown in §16. Since all the other indices in the formulas (1) and (2) are then invariants of the field K/k, it follows that the correction factor ϕ is an invari-

ant, too. In other words, ϕ is independent of the particular way in which the Herbrand unit group is chosen.

The invariants of K/k which we have connected here are, roughly speaking, of four types: well recognized invariants, such as Γ', ρ_1, ρ', n^* (in essence); invariants depending on the number of certain group extensions of groups belonging to K/k; invariants having to do with crossed characters; the invariant ϕ depending upon the explicit structure of E. In particular, all these invariants can be defined without reference to the auxiliary notion of factor set, except in the case of the invariant ϕ.

One may also specialize the composite formula to the cyclic case. The invariant n^* is a divisor of the order n and also a multiple of the order n, so that $n^* = n$. The correction factor ϕ is 1 according to Theorem 19.1. The order of the multiplicator $\prod_{i<j}(m_i, m_j)$ is always 1.

THEOREM 24.3. *If Z/k is a cyclic extension of k with the class group $\mathfrak{A}^t$, then the number $[F\mathfrak{A}^t : T\mathfrak{A}^t]$ of group extensions of this class group by the Galois group is related to the number of crossed characters of the Galois group in $\mathfrak{A}^t$ by the equation*

$$\omega(\mathfrak{A}', (A'))\,[U\mathfrak{A}^t : \mathfrak{A}^{t1-\sigma}] = \omega(A', E)\,[F\mathfrak{A}^t : T\mathfrak{A}^t]. \tag{4}$$

If the cyclic field and all its subfields have class number 1, all the terms in the equation (4) turn out to be 1.

25. **Examples.** As an illustration and check of our composite formula we consider in this section some examples of biquadratic fields. We take biquadratic fields over the field of all rational numbers which have, with all their quadratic subfields, the class number 1.

The field $K = R(7^{1/2}, i)$, with $i = (-1)^{1/2}$, is the join of two quadratic fields, $R(i)$ and $R(7^{1/2})$, each of which is known[58] to have class number 1. By Hilbert's form of a theorem of Dirichlet [27, p. 51] the class number of K is a factor of the product of the class numbers of $R(i)$ and $R(7^{1/2})$. Hence in K every ideal is principal. The same is true for the third quadratic subfield $R((-7)^{1/2})$ of K. The terms involving the deficiency ω and the terms involving ideal classes $\mathfrak{A}^t$ in the final formula for the abelian case are there all 1, hence formula (3) of §24 becomes

$$J = n^*(J(E)J(E^*)^{-1})\left[\prod_{i<j}(m_i, m_j)\right]^{-1}. \tag{1}$$

The Galois group of K/R is the four group, so the least common multiple J of the orders of the group elements is 2. The order $\prod_{i<j}(m_i, m_j)$ of the multiplicator turns out to be 2. The quantity n^* can be computed, as in §13, in terms of the ramification orders of the various rational primes in K. The three

[58] See the list of quadratic fields in the back of [35].

quadratic subfields $R(7^{1/2})$, $R(i)$, $R((-7)^{1/2})$ have respectively the discriminants $7\cdot 2^2$, -4, and 7. It follows then that the only ramified prime divisors for K/R are 2, 7, and p_∞, the infinite prime divisor of R. The ramification orders are $e_1=e_7=e_\infty=2$. All the ramification orders are thus divisors of the index J, so that, according to Theorem 13.2, $n^*=n=4$.

Finally, the correction factor $J(E)J(E^*)^{-1}$ may be found directly, by the normalization method of §19, if one first computes the group E of units in K. In the field $R(7^{1/2})$ the basic unit is $\rho=8+3(7^{1/2})$. By a computation resembling that in Lemma 3 of §19 one finally discovers that ρ is the relative norm of a unit $\eta=(3+7^{1/2})(1+i)2^{-1}$ in the field K. One then shows that every unit in K has the form $i^a\eta^b$. In the Galois group of K/k let β, γ, and δ denote respectively the automorphisms which leave fixed the subfield $R(i)$, $R(7^{1/2})$, and $R((-7)^{1/2})$. The generating units i and η then behave under the various automorphisms as follows:

$$i^\beta = i, \qquad i^\gamma = i^3 = i^{-1},$$
$$\eta^\beta = i\eta^{-1}, \qquad \eta^\gamma = i^2\eta, \qquad \eta^\delta = \eta^{-1}.$$

As an Herbrand subgroup one may choose the group generated by the unit $P=i^3\eta^2$, which has the appropriate behaviour, $P^\beta=P$, $P^{1+\gamma}=1$, relative to the generator β of the decomposition group $\{1, \beta\}$ belonging to the only infinite prime divisor of R. The correction factor then involves the groups D and D^* defined in Lemma 2 of §19. One finds that D is the group generated by i and η^2, while $D^*=E^*$. Therefore

$$J(E)J(E^*)^{-1} = [E:E^*][D:D^*]^{-1}[e:1]^{-1} = 8(4.2)^{-1} = 1.$$

These values, substituted in (1), give an identity.

A similar check has been carried out for the field $R(2^{1/2}, i)$. The ramifications are $e_\infty=2$, $e_2=4$, consequently $n^*=2$. The class numbers for this field and its quadratic subfields are again all 1, so that (1) holds. The group of units E is generated by a primitive 8th root of unity $\lambda=(1+i)(2^{1/2})^{-1}$ and by $\rho=1+2^{1/2}$. The unit ρ^2 may serve as generator of the Herbrand subgroup E^*. One finds $D^*=E^*$, while D is generated by i and ρ^2, whence the correction factor $J(E)J(E^*)^{-1}$ is 2. This again checks equation (1).

Another biquadratic field with class number 1 and with quadratic subfields of class number 1 is $R(3^{1/2}, i)$. The equation (1) again holds, this time with $e_2=e_3=e_\infty=2$, $n^*=4$, $E=\{\lambda, (1+3^{1/2})(1+i)2^{-1}\}$, λ a primitive 12th root of unity, and $J(E)J(E^*)^{-1}=1$.

Chapter V. Generalization of the classical symbols

26. **A conductor for factor sets.** One can construct an analogue to the conductor of the classical theory if one recalls that the conductor of a cyclic field Z/k can be defined in terms of the local conductors (see [37]). For a local

cyclic extension Z_P/k_p the conductor $c(Z_P/k_p)=p^h$ is defined as the smallest power of p such that $a\equiv 1 \pmod{p^h}$ implies that a is a norm.

We first prove the existence of a power of P which will have the analogous property for factor sets. Consider a normal extension K_P/k_p of degree $m_p=m$ with Galois group $\Delta(P)=\Delta$. Denote by $A_P=A$ the multiplicative group of the field K_P. Let $A_{\tau,\eta}$ denote the components of the factor sets in FA, while TA are the analogous transformation quantities[59].

THEOREM 26.1. *For any normal extension K_P/k_p there is a non-negative power P^h of the prime divisor P of K_P such that every factor set FA with components $F_{\tau,\eta}\equiv 1 \pmod{P^h}$ is a transformation set; in other words,*

$$F_{\tau,\eta}\equiv 1 \pmod{P^h} \text{ implies } F_{\tau,\eta}=TB_\tau \qquad \text{for suitable } B_\tau \text{ in } A. \tag{1}$$

Proof. Consider first the case where P is a finite prime divisor, so that (1) is an ordinary congruence modulo the ideal P^h. As in the cyclic case, we may then furnish a proof by using the P-adic logarithm, $\log A$, and the corresponding exponential, $\exp[(\log A)/m]$. The exponent h may be chosen so large that these expressions are uniquely defined[60] whenever $A\equiv 1 \pmod{P^h}$. Each such A then has in K_P a unique mth root, $B=\exp[(\log A)/m]$. For any factor set $F_{\tau,\eta}$ of (1) there will thus be a set of mth roots

$$B_{\tau,\eta}=\exp[(\log F_{\tau,\eta})/m]. \tag{2}$$

The associativity conditions for this set can then be derived from the conditions for $F_{\tau,\eta}$ by elementary properties of exponent and logarithm. Consequently $F_{\tau,\eta}=B_{\tau,\eta}^m$ is the mth power of a factor set, so it is a transformation set[61] by Lemma 1 of §1.

It remains to establish (1) in the case when P is an infinite prime divisor. To insure the validity of the theorem in this case, we must use a new convention as to congruences modulo such a divisor. An infinite prime divisor p_∞ of k is in effect an isomorphic map $\psi_p(a)=a'$ of k in a subfield k' of the field of all complex numbers. The valuation belonging to p_∞ is the valuation $\|a\|=|a'|$ induced on k by the natural valuation of k'. If p_∞ is real[62], then

$$a\equiv 1 \pmod{p_\infty} \qquad \text{if and only if } \psi_p(a)>0. \tag{3}$$

On the other hand, if p_∞ is complex[63], we define

---

[59] We omit the subscripts P. Of course, all terms are defined with respect to P.

[60] According to [23, p. 149], it suffices to take $h\geqq[s/(p_0-1)]+1+t$, where p_0 is the rational belonging to p and where $V_P(p_0)=s$, $V_P(m)=t$ in the P-adic valuation V_P of K_P, normalized by $V_P(P)=1$. Logarithm and exponential are also defined by Chevalley [10].

[61] This proof is a slight variant of the one given by Tannaka [38, Theorem 2]. We do not follow his exposition because in his conductor in the large he has omitted the (essential) infinite prime divisors.

[62] In other words, if k' consists of real numbers.

[63] In other words, if k' contains complex numbers.

(4) $a \equiv 1 \pmod{p_\infty}$ if and only if $a \neq 0$,

(5) $a \equiv 1 \pmod{p_\infty^2}$ if and only if $\psi_p(a)$ is real and positive.

A congruence modulo p_∞^h, $h>2$, is to have the same meaning as the same congruence for p_∞^2. This congruence (5) is an acceptable definition, because the elements $\equiv 1 \pmod{p_\infty^2}$ do form a multiplicative subgroup of the group of all nonzero elements of k.

These conventions differ at (5) from the usual ones (see [20, 22]). Our reasons for the adoption of such a definition are: first, we need a congruence modulo p_∞^2 which is not trivially true for all $a \neq 0$; secondly, the convention (5) agrees with the convention as to the ramifications. For, let k_∞ be the field of real numbers, K_∞ the field of complex numbers. Then the prime divisor p_∞[44] of k_∞ is ramified in K_∞, with $p_\infty = P_\infty^2$, according to the usual convention. The possible congruences then agree for any a in k_∞:

(6) $a \equiv 1 \pmod{p_\infty}$ if and only if $a \equiv 1 \pmod{P_\infty^2}$.

In other words, the formal decomposition $p_\infty = P_\infty^2$ may be actually substituted in congruences. This possibility of substitution may be readily extended to the decomposition of any p_∞ in a normal extension K/k, which would not be possible with the conventions of Hasse.

The infinite case of Theorem 26.1 is embodied in the following:

THEOREM 26.2. *Let $A_{\sigma,\tau}$ be a factor set of complex numbers belonging to the (cyclic) Galois group of the field K_∞ of complex numbers over the field k_∞ of real numbers. Then $A_{\sigma,\tau} \equiv 1 \pmod{P_\infty^2}$ implies that $A_{\sigma,\tau}$ is a transformation set from K_∞, provided this congruence is interpreted as in* (5).

Proof. The proof will depend on the analysis of the algebra determined by $A_{\sigma,\tau}$ as a quaternion algebra. The Galois group Δ of K_∞/k_∞ is a cyclic group of two elements $\{1, \zeta\}$. Consequently the crossed product $(K_\infty/k_\infty, \Delta, A_{\sigma,\tau})$ has a cyclic generation $(K_\infty/k_\infty, \zeta, a)$ where the constant a in the normalized generation may be computed as $a = A_{\zeta,\zeta}A_{1,1}$. The condition $A_{\zeta,\zeta} \equiv A_{1,1} \equiv 1 \pmod{P_\infty^2}$ means that a is real and positive, hence a norm in K_∞/k_∞. The algebra is therefore a total matric algebra, as asserted in Theorem 26.2.

The local conductor $C_P(K/k)$ for any P of K may be defined as the least power P^h, with $h \geqq 0$, for which the conclusion (1) of Theorem 26.1 will hold for the corresponding local extension. The conjugate prime divisors P^σ of P determine equivalent extensions, thus

(7) $$[C_P(K/k)]^\sigma = C_{P^\sigma}(K/k).$$

The characteristic property of these local conductors can be stated "in the large" as follows:

(44) That is to say, the natural valuation.

THEOREM 26.3. *Let F, G be two factor sets for Γ in K, such that at every prime divisor p of k one has for at least one factor P of p the multiplicative congruences*

$$F_{\xi,\eta} \equiv G_{\xi,\eta} \pmod{C_P(K/k)}; \qquad \xi, \eta \text{ in } \Delta(P). \tag{8}$$

Then one can conclude $F \sim G$; that is, that there is a transformation set TA with

$$F = GTA.$$

Proof. The hypothesis, as a multiplicative congruence, means that the quotient factor set FG^{-1} satisfies the conditions $(FG^{-1}) \equiv 1 \pmod{C_P(K/k)}$, for all P. According to formula (7) of §2 this factor set $(FG^{-1})_{\xi,\eta}$ is exactly one giving the local component $(K, \Gamma, FG^{-1})_p$ of the algebra determined by FG^{-1}. The congruences (1) therefore just suffice to insure that every local component is similar to one, which in turn implies $FG^{-1} \sim 1$ in the large, by the fundamental theorem for algebras over algebraic number fields.

As a matter of fact, the congruences (8) in this theorem need only be assumed for a finite number of prime divisors P.

THEOREM 26.4. *The local conductor $C_P(K/k)$ is 1 if and only if P is unramified in K/k.*

Proof. This conclusion is trivial for an infinite prime divisor P. For a finite P we recall the convention ([20, 23])

$$A \equiv 1 \pmod{P^0} \qquad \text{if and only if } A \text{ is relatively prime to } P. \tag{9}$$

Suppose first that P is unramified in K/k. Then the requirement $F_{\xi,\eta} \equiv 1 \pmod{P^0}$ means that $F_{\xi,\eta}$ is relatively prime to P. But by Theorem 12.1 a factor set of P-adic units is always similar to 1 for an unramified local extension. Hence P^0 is the local conductor. Conversely, suppose that the power P^0 functions as the local conductor. This means that every factor set of units is similar to 1. According to Theorem 12.1 the factor sets of local units yield a group of algebras with order equal to the ramification order e_p. Therefore they can all be similar to 1 only if $e_p = 1$.

This theorem suggests that we introduce as the conductor for factor sets of K/k the finite divisor

$$C(K/k) = \prod_P C_P(K/k). \tag{10}$$

This conductor need not agree with the ordinary conductor for abelian (noncyclic) fields. We remark that our principal class (see §6) $T\mathfrak{A}'(F'')$ contains the "ray" modulo $C(K/k)$. This is a strict generalization of the classical results, provided we take for the module M a multiple of $C(K/k)$.

Theorem 12.1 can also be used to obtain a somewhat closer estimate of the size of the conductor. According to this theorem, the eth power of any

factor set of local units is similar to 1. To make $F \sim 1$ we may then, as in (2), represent F as the eth power of a set: $B_{\xi,\eta} = \exp[(\log F_{\xi,\eta})/e_p]$; $\xi, \eta \; \Delta(P)$. The original congruence $F_{\xi,\eta} \equiv 1 \pmod{P^h}$ must then have a module h so chosen as to make the exponential and logarithmic series converge. Let V_P be the P-adic valuation of K, and assume

$$V_P(e_p) = r, \qquad V_P(p_0) = s,$$

p_0 the rational prime which is divisible by P.

Then, if $h > s/(p_0-1)$, $\log F$ converges and has P-adic order h. The exponential function has as argument a quantity $(\log F)/e_p$ of P-adic order $h-r$. To insure the convergence of the exponential series this must again exceed (see [23, p. 149]) $s/(p_0-1)$. Therefore $h > r+s/(p_0-1)$ suffices.

THEOREM 26.5. *If a finite prime divisor P is ramified in K/k in such a fashion that P^r is the highest power dividing the ramification order e_p, while P^s is the highest power dividing the rational prime p_0, then the local conductor $C_P(K/k)$ is a divisor of P^h, where $h = r+1+[s/(p_0-1)]$.*

COROLLARY. *If a finite prime divisor P has no higher ramifications in a normal extension K/R of the field of rational numbers, and if P is ramified in K/R and relatively prime to 2, then the local conductor $C_P(K/k)$ is P.*

In the case of a cyclic extension the factor sets can be normalized to single quantities. It follows that the conductor for factor sets is identical with the ordinary conductor([65]).

27. **A generalized norm residue symbol.** The investigations of Chevalley and Hasse [9, 11, 21, 22] showed that the norm residue symbols $((a, Z)/p)$ for a cyclic field Z/k can be defined purely locally, in terms of the invariants of the corresponding algebra (Z, λ, a). In a similar fashion, crossed products may be used to define a generalized "*norm residue*" symbol for a factor set F of a normal extension([66]) K/k. Unlike the classical symbol, this new symbol is not an element of the Galois group Γ of K/k, but its formal properties are similar to the ordinary ones (see [20, Part II]).

Let F be a factor set for K/k, p a prime divisor of k, and define

$$(*27.1) \quad (K/k, F; p) = \exp(\mu_p/n) = e^{2\pi i \mu_p/n},$$

where

$$\mu_p = \mu_p(K/k, \Gamma, F) \equiv \mu(K/k, \Gamma, F)_p \pmod{n}$$

is the additive invariant, mod n, of the algebra determined by F. Immediately one establishes the following properties of this symbol:

([65]) [37, Theorem 3]. Here also other inequalities for the conductor are given.

([66]) The introduction of such a general symbol is originally a suggestion due to Hasse.

$$(K, F_1; \mathfrak{p})(K, F_2; \mathfrak{p}) = (K, F_1F_2; \mathfrak{p}), \tag{1}$$

$$(K, F; \mathfrak{p})^{-1} = (K, F^{-1}; \mathfrak{p}), \tag{2}$$

$$(K, TA; \mathfrak{p}) = 1. \tag{3}$$

Hence the norm residue symbol maps the group FA/TA of similarity classes of factor sets homomorphically in the cyclic group of nth roots of unity.

To further analyze (3), recall that the $\mathfrak{p}$-component of a crossed product (see §2, (7)) is $(K, \Gamma, F)_{\mathfrak{p}} \sim (K_P, \Delta(P), F \cap \Delta)$ where $F \cap \Delta$ is the factor set consisting of those elements $F_{\zeta,\eta}$ of F whose indices ζ, η lie in the decomposition group $\Delta = \Delta(P)$ of a fixed factor P of $\mathfrak{p}$. Hence

$$(K, F; \mathfrak{p}) = 1 \qquad \text{if and only if } F \cap \Delta \sim 1. \tag{4}$$

In terms of the local conductor C_P, this may be written as

$$(K, F; \mathfrak{p}) = 1 \qquad \text{if and only if } (F \cap \Delta)(TA_P)^{-1} \equiv 1 \pmod{C_P} \tag{5}$$

for a suitable transformation set[67] TA_P. In terms of multiplicative congruences (see [10, 20, 23]), this last congruence can be written as $F_{\zeta,\eta} \equiv (A_P)_\zeta (A_P)_\eta^\zeta (A_P)_{\zeta\eta}^{-1} \pmod{C_P}$. This result specifies the sense in which our symbol is a "norm residue" symbol. Thus we have proved

THEOREM 27.1. *The symbol $(K, F; \mathfrak{p})$ depends only on the (multiplicative) residue classes of the elements $F_{\zeta,\eta}$ modulo the conductor C_P, where ζ, η run over the elements in the decomposition group of some prime factor P of $\mathfrak{p}$ in K.*

THEOREM 27.2. *If $(K, F; \mathfrak{p}) = 1$ for every factor set F, then $\mathfrak{p}$ is totally decomposed in K; that is, $\mathfrak{p}$ has in K n distinct prime factors $\mathfrak{p} = P_1P_2 \cdots P_n$.*

Proof. If P is not totally decomposed in K, then each factor P of $\mathfrak{p}$ has a decomposition group $\Delta(P)$ of order $[K_P : k_\mathfrak{p}] = m_\mathfrak{p}$ greater than 1. If l is any rational prime factor of the degree $m_\mathfrak{p}$, there must then be a local algebra $S_\mathfrak{p}$ which has the invariant $\mu(S_\mathfrak{p}) = n/l \not\equiv 0 \pmod{n}$. By Theorem 14.1 the algebra $S_\mathfrak{p}$ is the component of an algebra S in the large. Any factor set F for such an algebra S would have $(K, F; \mathfrak{p}) \neq 1$, counter to the assumption of the theorem. The sum relation[68] for the invariants of an algebra implies a product formula for the norm residue symbol

$$\prod_{\mathfrak{p}} (K, F; \mathfrak{p}) = 1. \tag{6}$$

The question of the existence of a factor set F with certain given values for $(K, F; \mathfrak{p})$ is identical with the question (see §§13–14) of the existence of actual algebras S with specified local components.

Consider next a normal subfield L/k of K, belonging to the subgroup Λ

[67] The associated vector $(A_P)_\zeta$ has components in K_P for ζ in Δ.

[68] As stated in the proof of Theorem 6.1.

of Γ. Let $F_\Lambda = F'$ be a factor set belonging to L/k. A formula of the theory of algebras (see [13, p. 63]) expresses the crossed product given by F' as a crossed product for K/k by the relation

$$(L/k, \Gamma/\Lambda, F') \sim (K/k, \Gamma, F' \cup \Gamma), \tag{7}$$

where $F' \cup \Gamma$ is the factor set obtained by extension of F' with Γ, as in

$$F' \cup \Gamma = \{F_{\sigma,\tau}\} \qquad \text{where } F_{\sigma,\tau} = F'_{\sigma\Lambda,\tau\Lambda} = F'_{\bar{\sigma},\bar{\tau}}, \tag{8}$$

for $\bar{\sigma}$, $\bar{\tau}$ in Γ/Λ and σ, τ arbitrary representatives of the cosets $\bar{\sigma}$, $\bar{\tau}$. The factor set $F' \cup \Gamma$ is "Λ-*symmetric*." In general, a factor set $F_{\sigma,\tau}$ is called Λ-symmetric if the components $F_{\sigma,\tau}$ depend only on the position of σ and τ in the cosets of Γ/Λ.

Since similar algebras have the same invariants, (8) yields

$$(L/k, F'; \mathfrak{p}) = (K/k, F' \cup \Gamma; \mathfrak{p}). \tag{9}$$

This rule determines $(K, F; \mathfrak{p})$ only for those factor sets F which are similar to Λ-symmetric factor sets(59). Rule (9) is the analogue of the simple rule "$((a, K)/\mathfrak{p})$ induces $((a, L)/\mathfrak{p})$" of the classical theory.

Consider next an arbitrary finite extension L of the base field k, which "translates" the extension K/k to the extension KL/L. The Galois group of the join KL/L is isomorphic to that subgroup Λ of Γ which leaves fixed all elements of the intersection $K \cap L$. On a crossed product the extension of the base field k to L has the effect (see [13, p. 61]):

$$(K/k, \Gamma, F)_L \sim (KL/L, \Lambda, F \cap \Lambda), \tag{10}$$

where $F \cap \Lambda$ is the factor set whose components are those elements $F_{\sigma,\tau}$ of F which have subscripts σ and τ in the group Λ. On the other hand (see [28, Theorem 5]), it can be shown that for any prime divisor q in L the q-invariant of an extended algebra S_L is obtained by multiplying the additive $\mathfrak{p}$-invariant (mod 1) of S by $h = [L_q : k_\mathfrak{p}]$, where $q \mid \mathfrak{p}$. In terms of the norm residue symbol, these two facts may be combined as the "*translation law*"

$$(K/k, F; \mathfrak{p})^h = (KL/L, F \cap \Lambda; q). \tag{11}$$

A special case of (11) is the rule for a subfield: if $k \subset L \subset K$, while $P' \mid \mathfrak{p}$ is a prime divisor in L, Λ the corresponding group, then

$$(K/L, F \cap \Lambda; P') = (K/k, F; \mathfrak{p})^h \tag{12}$$

where $h = [L_{P'} : k_\mathfrak{p}]$.

Consider a field K which is the join over k of two fields K_1 and K_2, each normal over k. In the Galois group Γ of K/k each automorphism σ induces

(59) Even if F is not Λ-symmetric, its local component $F \cap \Delta$ at some of the $P \mid \mathfrak{p}$ may be so symmetric; this makes possible a slight extension of (9).

an automorphism σ_1 of K_1/k and an automorphism σ_2 of K_2/k. Since K is generated by K_1 and K_2, the original σ is uniquely determined by its homomorphic maps σ_1 and σ_2, so σ may be represented by the formal product([70]) $\sigma=\sigma_1\sigma_2$. As in (7) a crossed product $(K_1/k, \Gamma_1, F_1)$ for the first extension is similar to a crossed product $(K_1K_2/k, \Gamma, \overline{F})$, where the factor set F is obtained from F_1 by the formulas (8) as

$$\overline{F}_{\sigma_1\sigma_2,\tau_1\tau_2} = F_{1^{\sigma_1},\tau_1}, \qquad \sigma_1\sigma_2 \text{ and } \tau_1\tau_2 \text{ in } \Gamma. \tag{13}$$

The crossed products $(K_2/k, \Gamma_2, F_2)$ are similarly extended. For the direct product of two such algebras, we then have

$$(K_1/k, \Gamma_1, F_1) \times (K_2/k, \Gamma_2, F_2) \sim (K_1K_2/k, \Gamma, \overline{F}) \tag{14}$$

where the factor set $\overline{F}$ is given by formulas such as (13)

$$\overline{F}_{\sigma_1\sigma_2,\tau_1\tau_2} = F_{1^{\sigma_1},\tau_1}F_{2^{\sigma_2},\tau_2} \qquad \text{with } \sigma_1 \text{ and } \tau_1 \text{ in } \Gamma_1,\ \sigma_2 \text{ and } \tau_2 \text{ in } \Gamma_2.$$

This means that the matrix of $\overline{F}$ is obtained from the *Kronecker product* $F_1 \otimes F_2$ of the given matrices by taking that submatrix belonging to the group Γ (see [34, p. 691]). This yields for the norm residue symbols a rule

$$(K_1K_2/k, (F_1 \otimes F_2) \cap \Gamma; p) = (K_1/k, F_1; p)(K_2/k, F_2; p). \tag{15}$$

Formula (15) is a direct generalization of one of the rules for the ordinary norm residue symbol (see [20, Part II, p. 27, (8)]).

The rules for the values taken on by the classical norm residue symbol (see [20, Part II, p. 35, VI and VII]) break down in our case. The p-invariant of an algebra S always has the form $x(n/m_p)$ for an integer x. Hence the corresponding norm residue symbol is an m_pth root of unity. Not all such roots need occur, because not all local algebras are components of algebras in the large (see §14).

THEOREM 27.3. *If p is unramified in K/k, the norm residue symbol $(K/k, F; p)$ for variable factor sets F takes on all values* exp $(1/m_p)$ *in a cyclic group of order m_p. If p is ramified, the norm residue symbol takes on some, but not necessarily all, of these values.*

For factor sets F which are relatively prime to a given prime divisor p the corresponding local factor set consists of P-adic units, and the results of Theorems 14.4 and 14.5 determine the range of values of the norm residue symbol in this case.

THEOREM 27.4. *If p is unramified in K/k, the norm residue symbol $(K/k, \Gamma; p)$ for variable factor sets F which are relatively prime to p is always* 1. *If p is ramified, the norm residue symbol for F relatively prime to p takes on all*

([70]) Here σ_1 is an automorphism of K_1 (not of K). This formula represents Γ as a subgroup of the direct product $\Gamma_1 \times \Gamma_2$ of the groups Γ_1 and Γ_2 of K_1/k, K_2/k, respectively.

values exp $(y/(e_p, \mathfrak{n}))$, *with* y *integral, in a cyclic group of order* $(e_p, \mathfrak{n})$. *Here* $\mathfrak{n}=\mathfrak{n}_i$ *is determined from* $p=p_i$ *as in Theorem* 14.5.

28. **A generalized Artin symbol.** The invariants of ideal factor sets, as introduced in §4, may be used to set up an Artin symbol resembling the symbol $(K/\mathfrak{a})$ of the classical theory ([20, Part II]). The ideal $\mathfrak{a}$ which is relatively prime to the conductor is to be replaced by a factor set $\mathfrak{F}$ of ideals which are relatively prime to the conductor $C(K/k)$. We define[71]

$$*(28.1)\quad (K\,|\,\mathfrak{F})=(K/k;\mathfrak{F})=\exp\left\{\textstyle\sum_p \mu_p(\mathfrak{F})n^{-1}\right\},$$

where the sum is to be taken over all unramified prime divisors p for K/k, while μ_p is to be the invariant defined in (*5.6). In terms of the explicit formula §5, (16) for this invariant we may also write, for any factor P of p,

$$(1)\qquad (K\,|\,\mathfrak{F}) = \exp\left\{\sum_p \sum_{j=0}^{m-1} m^{-1}V_P(\mathfrak{F}^{\delta^j,\mathfrak{a}})\right\}$$

where $V_P(\mathfrak{A})$ denotes the P-adic order of an ideal $\mathfrak{A}$, while $\delta=[(K/k)/P]$ is the Frobenius automorphism for P, of order $m=m_p$.

Our symbol may be obtained for any $\mathfrak{F}$ once its values are known for the p-primary factor sets. We know that the invariant $\mu_p(\mathfrak{F})$ depends solely on the p-primary factor $\mathfrak{F}^{(p)}$ of $\mathfrak{F}$. Specifically, any $\mathfrak{F}$ which is relatively prime to the conductor can be represented uniquely by a product

$$(2)\qquad \mathfrak{F} = \mathfrak{F}^{(1)}\mathfrak{F}^{(2)}\cdots\mathfrak{F}^{(s)},$$

where $\mathfrak{F}^{(i)}$ is a p_i-primary factor set, while $p_1, p_2, \cdots, p_s$ are all the various (unramified) prime divisors of k involved in $\mathfrak{F}$. Since $\mu_i(\mathfrak{F})=\mu_i(\mathfrak{F}^{(i)})$, we have

$$(3)\qquad (K\,|\,\mathfrak{F}) = (K\,|\,\mathfrak{F}^{(1)})(K\,|\,\mathfrak{F}^{(2)})\cdots(K\,|\,\mathfrak{F}^{(s)}).$$

For a p-primary factor set $\mathfrak{F}^{(p)}$ all invariants $\mu_q(\mathfrak{F}^{(p)})$ for $q\neq p$ are zero, so the original definition of the symbol may be rephrased as

$$(4)\qquad (K\,|\,\mathfrak{F}^{(p)}) = \exp\left\{\mu_p(\mathfrak{F}^{(p)})/n\right\}.$$

The equations (4), (3), (2) might have been adopted to define the Artin symbol $(K\,|\,\mathfrak{F})$. Alternatively, (4) may be inverted to give a definition of the p-invariants in terms of the Artin symbol:

$$(5)\qquad \mu_p(\mathfrak{F}) = (n/2\pi i)\log\left[(K\,|\,\mathfrak{F}^{(p)})\right],$$

where $\mathfrak{F}^{(p)}$ denotes the p-primary factor of the given ideal factor set $\mathfrak{F}$.

THEOREM 28.1. *The Artin symbol affords an isomorphic mapping* $\mathfrak{F}\to(K\,|\,\mathfrak{F})$ *of the group* $F\mathfrak{A}'/T\mathfrak{A}'(F'')$ *onto the group of* J*th roots of unity.*

[71] We let $\exp u = e^{2\pi i u}$.

The quotient group in question is exactly the group whose order was computed in Theorem 6.1. Much as in Artin's reciprocity theorem, the Artin symbol thus furnishes an explicit realization of the isomorphism of the class group $F\mathfrak{A}'/T\mathfrak{A}'(F'')$ to a substitute for the Galois group (Jth roots of unity).

Proof. To show that $\mathfrak{F}\to(K\,|\,\mathfrak{F})$ is a homomorphism, we need only refer to the corresponding properties of the invariant $\mu_p(\mathfrak{F})$, as given in §5. We find

$$(K\,|\,\mathfrak{F}_1)(K\,|\,\mathfrak{F}_2) = (K\,|\,\mathfrak{F}_1\mathfrak{F}_2), \tag{6}$$

$$(K\,|\,\mathfrak{F})^{-1} = (K\,|\,\mathfrak{F}^{-1}), \tag{7}$$

$$\mathfrak{F}_1 \sim \mathfrak{F}_2 \text{ implies } (K\,|\,\mathfrak{F}_1) = (K\,|\,\mathfrak{F}_2). \tag{8}$$

Lemma 5.1, combined with the Tschebotareff density theorem, indicates that every Jth root of unity appears as the value of some Artin symbol $(K\,|\,\mathfrak{F})$. If $\mathfrak{F}$ is taken as the principal ideal (F''), then by Theorem 6.1 $\mathfrak{F}$ has the same invariants as the algebra $(K/k, \Gamma, F'')$. The sum of these invariants is 0 (mod n), hence

$$\mathfrak{F} \text{ in } (F'') \text{ implies } (K\,|\,\mathfrak{F}) = 1. \tag{9}$$

It remains only to prove

$$(K\,|\,\mathfrak{F}) = 1 \text{ implies } \mathfrak{F} \text{ lies in } T\mathfrak{A}'(F''). \tag{10}$$

But $(K\,|\,\mathfrak{F}) = 1$ means according to definition (*28.1) that $\sum_p \mu_p(\mathfrak{F}) \equiv 0 \pmod{n}$. This implies that the invariants $\mu_p(\mathfrak{F})$ are the invariants of an actual algebra S' relatively prime to M. By Theorem 6.2, S' has a crossed product representation with a factor set (F''). By Theorem 5.4, $\mu_p((F'')) \equiv \mu_p(S') \equiv \mu_p(\mathfrak{F})$, so that, by Theorem 5.2, $\mathfrak{F}(F'')^{-1}$ is a transformation set $T\mathfrak{A}'$, as asserted.

Since the norm residue symbol can also be defined (see §27) in terms of invariants of algebras, we may obtain it from the Artin symbol, applied to the p-primary component $(F)^{(p)}$ of a factor set of principal ideals

$$(K, F; p) = (K\,|\,(F)^{(p)}), \quad \text{if } p \text{ is unramified in } K/k. \tag{11}$$

According to Lemma 4 of §5 the p-invariants $\mu_p((F)^{(p)})$ run over all multiples of nm_p^{-1}. Hence we obtain the following result, which resembles the determination of the decomposition of unramified prime divisors in terms of the abelian class groups (see [20, Part II]).

THEOREM 28.2. *If p is unramified in K/k the degree $f = m_p$ of a prime factor of p is the least integer f such that $(K\,|\,\mathfrak{F}^{(p)})^f = 1$ for every p-primary factor set $\mathfrak{F}^{(p)}$.*

Remark. In general not every coset of $F\mathfrak{A}'/T\mathfrak{A}'(F'')$ will contain p-primary factor sets. The least common multiple $J(\Gamma)$ may be a proper multiple of every residue class degree f which is permitted by the structure of the Galois group. The simplest example is furnished by fields K/k with the sym-

metric group of 3 letters as the Galois group. Here $J(\Gamma)=2,3$ and $f=2,3$. Thus, the two classes of order 6 in the cyclic group $F\mathfrak{A}'/T\mathfrak{A}'(F'')$ are free from primary factor sets. This is another instance which illustrates the discrepancy between the classical class field theory and our theory.

The cosets which contain primary factor sets can easily be determined by means of Theorem 28.1. A coset contains a p-primary factor set if its order in the class group is equal to a permissible residue class degree of an unramified prime divisor. These degrees are exactly the orders of the cyclic subgroups of Γ, as follows from the ramification theory. Now $J(\Gamma)$ is the least common multiple of the orders of the group elements, hence $J(\Gamma)\equiv 0 \pmod f$ for every possible f. The Frobenius density theorem then proves that at least one class of order f must contain infinitely many primary factor sets.

Consider next the symbol referred to a subfield L of K such that

$$k \subset L \subset K, \qquad L/k \text{ normal.} \tag{12}$$

If Λ is the subgroup of Γ belonging to L, then the Galois group of L/k is Γ/Λ. If a prime divisor P of K has a multiple Q in L, then the Frobenius automorphism $[(K/k)/P]=\delta$ of P induces in Γ/Λ the Frobenius automorphism $[(L/k)/Q]=\delta\Lambda=\zeta$ of Q (see [20, Part II, p. 6, Rule III]). A factor set $\mathfrak{G}$ for Λ can be extended to a factor set $\mathfrak{F}$ for K by the rules

$$\mathfrak{F}_{\sigma,\tau} = \mathfrak{G}_{\sigma\Lambda,\tau\Lambda}; \tag{13}$$

the result is a Λ-symmetric factor set of K. By (1), the Artin symbol is

$$(K\,|\,\mathfrak{F}) = \exp\left\{\sum_p \sum_{i=0}^{m-1} m^{-1} V_P(\mathfrak{F}_{\tau^i,\xi})\right\}, \qquad m = m_p.$$

The order V_P is the same as V_Q, since p is unramified, while the order n_p of the automorphism ζ is a divisor $n_p=m_p/r$ of the order m_p of δ, so this equation becomes

$$(K\,|\,\mathfrak{F}) = \exp\left\{\sum_p (r/m_p) \sum_i V_Q(\mathfrak{G}_{\xi^i,\xi})\right\}, \qquad i = 0, 1, \cdots, n_p - 1.$$

Since $r/m_p=1/n_p$, this yields the result

$$(K/k;\mathfrak{F}) = (L/k;\mathfrak{G}). \tag{14}$$

Given the situation (12), with $\mathfrak{F}$ defined by (13), this gives the Artin symbol $(K\,|\,\mathfrak{F})$ for the Λ-symmetric sets $\mathfrak{F}$.

On the other hand, one may consider the symbol for K/L.

THEOREM 28.3. *If the subgroup Λ of Γ corresponds to the subfield $L\subset K$, while $\mathfrak{F}'\cap\Lambda$ denotes that part of the factor set $\mathfrak{F}'$ which refers to the subgroup Λ, then*

$$(K/L; \mathfrak{F}' \cap \Lambda) = (K/k; \mathfrak{F}')^d, \qquad \text{where } d = [L:k]. \tag{15}$$

This is a special case of the following general "translation law" for the norm residue symbol:

THEOREM 28.4. *Let L be any finite extension of k such that the intersection $L \cap K$ belongs to the subgroup Λ of Γ. Then*

$$(KL/L; \mathfrak{F}' \cap \Lambda) = (K/k, \mathfrak{F}')^d \qquad \text{where } d = [L:k]. \tag{16}$$

Proof. It will suffice to prove (16) for the case of a p-primary factor set $\mathfrak{F} = \mathfrak{F}^{(p)}$, where p is unramified in K/k. Let the decomposition of p in L be

$$p = q_1^{e_1} q_2^{e_2} \cdots q_h^{e_h}, \qquad N_{L/k} q_j = p^{f_j} \qquad (j = 1, \cdots, h). \tag{17}$$

In the join KL choose for each j a factor Q_j of q_j, and let P_j be the prime factor of Q_j which lies in K. The translation rule for the ordinary Artin symbol asserts that q_j is unramified in KL/L and that the Frobenius automorphism δ_j of Q_j (see [20, Part II, p. 8]) is

$$\delta_j = [(KL/L)/Q_j] = [(K/k)/P_j]^{f_j}.$$

Since the original Frobenius automorphism $\delta = [(K/k)/P_j]$ has order m_p, the derived automorphism δ_j must have the order

$$n_j = m_p (m_p, f_j)^{-1}, \qquad \text{where } N_{KL/L} Q_j = q_j^{n_j}.$$

In L the only primary components of $\mathfrak{F}$ are the q_j-primary ones, for $j = 1, \cdots, h$. Hence the translated Artin symbol has by definition the value

$$(KL/L; \mathfrak{F} \cap \Lambda) = \exp \left\{ \sum_{j=1}^{h} n_j^{-1} \left(\sum_{i=0}^{n_j - 1} W_j[\mathfrak{F}_{\delta_j^i, \delta_j}] \right) \right\},$$

where W_j denotes the valuation belonging to Q_j. Since $P_j \mid p$ is unramified, the ramification order of $Q_j \mid P_j$ must be exactly the ramification order e_j which figures in the decomposition (17). Therefore the power $W_j[\mathfrak{A}]$ to which Q_j divides any ideal $\mathfrak{A}$ of K is e_j times the power $V_j[\mathfrak{A}]$ to which P_j divides the same ideal, and

$$(KL/L; \mathfrak{F} \cap \Lambda) = \exp \left\{ \sum_{j=1}^{h} (e_j/n_j) \left(\sum_{i=0}^{n_j - 1} V_j[\mathfrak{F}_{\delta^i, \delta}] \right) \right\}.$$

But the order V_j of the product involved here was computed in (18) of Lemma 5.7, and is $(n_j f_j / n) \mu_p(\mathfrak{F})$, where $r = n_j$ is the order of the cyclic group on δ_j, while $t = f_j$ is the exponent in $\delta_j = \delta^t$. Therefore

$$(KL/L; \mathfrak{F} \cap \Lambda) = \exp \left\{ \sum_{j=1}^{h} (e_j f_j / n) \mu_p(\mathfrak{F}) \right\}. \tag{18}$$

But $\sum e_i f_i$ is exactly the degree $d=[L:k]$, while $\exp\{\mu_p(\mathfrak{F})/n\}$ is exactly $(K|\mathfrak{F})$ by (4). Therefore (18) gives the desired conclusion (16) of the theorem.

This translation law seems superficially out of agreement with the analogous law (see [20, Part II]) for the norm residue symbol, because in that case the exponent was $[L_q:k_p]$, while here it is $[L:k]$. The divergence disappears if one recalls that the factor set $\mathfrak{F}^{(p)}$ splits up into h p-primary factor sets over L, each of which involves an exponent $[L_q:k_p]$, for a total $[L_q:k_p]h=[L:k]$. This remark may be explicitly checked by using the theorem of (11).

As in §27 we may consider the case of the join $K=K_1K_2$ of two normal fields, and a factor set $(\mathfrak{F}_1\otimes\mathfrak{F}_2)\cap\Gamma$ obtained from the Kronecker product of two ideal factor sets $\mathfrak{F}_1$ and $\mathfrak{F}_2$ which are relatively prime to the conductor of K/k. This factor set can be represented as a product(72)

$$[(\mathfrak{F}_1\otimes I_2)\cap\Gamma]\cdot[(I_1\otimes\mathfrak{F}_2)\cap\Gamma],$$

to each factor of which the subfield rule (14) for symmetric factor sets may be applied. The result is

$$(19)\qquad (K_1K_2|\,(\mathfrak{F}_1\otimes\mathfrak{F}_2)\cap\Gamma)=(K_1|\,\mathfrak{F}_1)\cdot(K_2|\,\mathfrak{F}_2).$$

Bibliography

A. A. Albert

1. *Modern Higher Algebra*, Chicago, 1937.
2. *Structure of Algebras*, American Mathematical Society Colloquium Publications, vol. 24, 1939.
3. (With H. Hasse.) *A determination of all normal division algebras over an algebraic number field*, these Transactions, vol. 34 (1932), pp. 722–730.

Y. Akizuki

4. *Eine homomorphe Zuordnung der Elemente der Galoisschen Gruppe zu den Elementen einer Untergruppe der Normklassengruppe*, Mathematische Annalen, vol. 112 (1936), pp. 566–571.

E. Artin

5. *Über Einheiten relativ-galoisscher Zahlkörper*, Journal für die reine und angewandte Mathematik, vol. 167 (1932), pp. 153–156.

R. Baer

6. *Erweiterungen von Gruppen und ihren Isomorphismen*, Mathematische Zeitschrift, vol. 38 (1934), pp. 375–416.
7. *Automorphismen von Erweiterungsgruppen*, Actualités Scientifiques et Industrielles, vol. 205, Paris, 1935.

R. Brauer, H. Hasse, and E. Noether

8. *Beweis eines Hauptsatzes in der Theorie der Algebren*, Journal für die reine und angewandte Mathematik, vol. 167 (1932), pp. 399–404.

C. Chevalley

9. *La théorie du symbole de restes normiques*, Journal für die reine und angewandte Mathematik, vol. 169 (1932), pp. 141–157.
10. *Sur la théorie du corps de classes pour les corps finis et les corps locaux*, Journal of the Faculty of Science, University of Tokyo, section 1, vol. 2 (1933), pp. 365–474.

(72) Here I_i denotes the unit matrix of $[K_i:k]$ rows and columns.

11. *La théorie du corps de classes*, Annals of Mathematics, vol. 41 (1940), pp. 394–418.

A. H. CLIFFORD AND S. MACLANE

12. *Factor sets of a group and its abstract unit group*, forthcoming in these Transactions.

M. DEURING

13. *Algebren*, Ergebnisse der Mathematik, vol. 4, Berlin, 1934.

L. E. DICKSON

14. *New division algebras*, these Transactions, vol. 28 (1926), pp. 207–234.

R. FRUCHT

15. *Über die Darstellungen endlicher abelscher Gruppen durch Kollineationen*, Journal für die reine und angewandte Mathematik, vol. 166 (1931), pp. 16–29.

W. GRUNWALD

16. *Ein algebraisches Existenztheorem für algebraische Zahlkörper*, Journal für die reine und angewandte Mathematik, vol. 169 (1933), pp. 103–107.

M. HALL

17. *Group rings and extensions* I, Annals of Mathematics, vol. 39 (1938), pp. 220–234.

H. HASSE AND K. HENSEL

18. *Über die Normenreste eines relativ-zyklischen Körpers vom Primzahlgrad l nach einem Primteiler* $\mathfrak{l}$ *von l*, Mathematische Annalen, vol. 90 (1923), pp. 262–278.

H. HASSE

19. *Zwei Existenztheoreme für algebraische Zahlkörper*, Mathematische Annalen, vol. 95 (1925), pp. 229–238.

20. *Bericht über neuere Untersuchungen und Probleme der algebraischen Zahlkörper*, Jahresbericht der Deutschen Mathematiker-Vereinigung, Supplementary Volume 6 (1930).

21. *Theory of cyclic algebras over an algebraic number field*, these Transactions, vol. 34 (1932), pp. 171–214.

22. *Die Struktur der R. Brauerschen Algebrenklassengruppe über einem algebraischen Zahlkörper*, Mathematische Annalen, vol. 107 (1933), pp. 731–760.

23. *Klassenkörpertheorie*, mimeographed lecture notes, Marburg, 1933.

K. HENSEL

24. *Die Verallgemeinerung des Legendreschen Symbols für allgemeine algebraische Körper*, Journal für die reine und angewandte Mathematik, vol. 147 (1916), pp. 233–248.

J. HERBRAND

25. *Nouvelle démonstration et généralisation d'un théorème de Minkowski*, Comptes Rendus de l'Académie des Sciences, Paris, vol. 191 (1930), p. 1282.

26. *Sur les unités d'un corps algébrique*, ibid., vol. 192 (1931), p. 24.

D. HILBERT

27. *Collected Papers*, vol. 1.

G. KÖTHE

28. *Erweiterung des Zentrums einfacher Algebren*, Mathematische Annalen, vol. 107 (1932), pp. 761–766.

S. MACLANE (AND O. F. G. SCHILLING)

29. *Normal algebraic number fields*, Proceedings of the National Academy of Sciences, vol. 26 (1940), pp. 122–125.

M. MORIYA

30. *Über die Klassenzahl eines relativ-zyklischen Zahlkörpers von Primzahlgrade*, Japanese Journal of Mathematics, vol. 10 (1933), pp. 1–18.

T. NAKAYAMA

31. *Über die Beziehungen zwischen den Faktorensystemen und der Normklassengruppe eines Galoisschen Erweiterungskörpers*, Mathematische Annalen, vol. 112 (1936), pp. 85–91.

E. NOETHER

32. *Hyperkomplexe Systeme und ihre Beziehungen zur kommutativen Algebra und Zahlentheorie*, Verhandlungen des Internationalen Kongresses, Zürich (1932), vol. 1, pp. 189–195.
33. *Der Hauptgeschlechtssatz für relativgaloissche Zahlkörper*, Mathematische Annalen, vol. 108 (1933), pp. 411–419.

O. F. G. SCHILLING

34. *A generalization of local class field theory*, American Journal of Mathematics, vol. 60 (1938), pp. 667–704.

J. SOMMER

35. *Vorlesungen über Zahlentheorie*, Berlin, 1907.

I. SCHUR

36. *Untersuchungen über die Darstellungen der endlichen Gruppen durch gebrochene lineare Substitutionen*, Journal für die reine und angewandte Mathematik, vol. 132 (1907), pp. 85–137.

T. TANNAKA

37. *Eine Bemerkung über den Normensatz relativ-Galoisscher Zahlkörper*, Proceedings of the Imperial Academy of Tokyo, vol. 9 (1933), pp. 565–567.
38. *Über eine Indexrelation*, Science Reports of the Tôhoku Imperial University, (1), vol. 23 (1934–1935), pp. 343–358.

A. WEIL

39. *Remarques sur les résultats récents de C. Chevalley*, Comptes Rendus de l'Académie des Sciences, Paris, vol. 203 (1936), p. 208.
40. *Zur algebraischen Theorie der algebraischen Funktionen*, Journal für die reine und angewandte Mathematik, vol. 179 (1938), pp. 129–133.

E. WITT

41. *Zwei Regeln über verschränkte Produkte*, Journal für die reine und angewandte Mathematik, vol. 173 (1935), pp. 191–192.

B. L. VAN DER WAERDEN

42. *Moderne Algebra*, Part I, 2d edition, Berlin, 1939.

H. ZASSENHAUS

43. *Lehrbuch der Gruppentheorie*, Leipzig, 1937.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.
UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

FACTOR-SETS OF A GROUP IN ITS ABSTRACT UNIT GROUP

BY

A. H. CLIFFORD AND SAUNDERS MacLANE

The following interesting problem in group extensions arose in an arithmetical study of normal algebraic number fields(1).

Let Γ be any finite group of order n with elements $\rho, \sigma, \tau, \cdots$. By the *abstract unit group* of Γ we mean the free abelian group $\mathfrak{H}$ of rank $n-1$ generated by n symbols H^σ (σ ranging over Γ) subject to the single condition that their product is 1:

$$\prod_\sigma H^\sigma = 1.$$

Every element of $\mathfrak{H}$ is expressible in the form $H^{\Sigma a(\sigma)\sigma}$ with integers $a(\sigma)$, and

$$H^{\Sigma a(\sigma)\sigma} = H^{\Sigma b(\sigma)\sigma}$$

if and only if there is an integer c such that

$$a(\sigma) - b(\sigma) = c, \qquad \text{all } \sigma \text{ in } \Gamma.$$

($\mathfrak{H}$ is isomorphic with the additive group of the integral group ring of Γ reduced modulo the element $\sum\sigma$.)

For each τ in Γ, the mapping $H^\sigma \to H^{\sigma\tau}$ generates an automorphism of $\mathfrak{H}$, namely,

$$H^{\Sigma a(\sigma)\sigma} \to H^{\Sigma_\sigma a(\sigma)\sigma\tau} = H^{\Sigma_\sigma a(\sigma\tau^{-1})\sigma},$$

and to the product $\tau\rho$ evidently corresponds the product of the corresponding automorphisms:

Presented to the Society, February 22, 1941; received by the editors August 24, 1940.

(1) S. MacLane and O. F. G. Schilling, *Normal algebraic number fields*, these Transactions, vol. 50 (1941), pp. 295–384. The present paper is entirely independent thereof.

To indicate briefly what the connection is, let K be a real algebraic field normal of degree n over the rational field k, and let Γ be its Galois group. A theorem of Minkowski (Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, 1900, p. 90) asserts the existence of a unit H in K such that the group $\mathfrak{H}$ generated by H and all its conjugates H^σ is a free abelian group of rank $n-1$, the only relation satisfied by the H^σ being that asserting that the norm of H is 1. This group $\mathfrak{H}$ is of finite index in the group of all units of K. MacLane and Schilling study the class field theory of K by introducing crossed products (K, Γ, F). (K, Γ, F) is a normal simple algebra over k, determined by K, Γ, and a factor-set F of elements $F_{\sigma,\tau} \neq 0$ of K satisfying (0.3). Under the homomorphism $F_{\sigma,\tau} \to (F_{\sigma,\tau})$ carrying numerical factor-sets into factor-sets of principal ideals, those mapped into the identity are precisely the factor-sets of units. By further devices they are enabled in some cases to reduce the consideration of the latter to that of factor-sets lying in the Minkowski subgroup $\mathfrak{H}$.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

$$A^{\tau\rho} = (A^{\tau})^{\rho}, \qquad \text{all } A \text{ in } \mathfrak{H}. \tag{0.1}$$

This representation of Γ by automorphisms of $\mathfrak{H}$ is effectively the right regular representation of Γ applied to the basis H^{σ} of $\mathfrak{H}$.

The problem in question is to determine all extensions(2) $\mathfrak{G}$ of $\mathfrak{H}$ by Γ realizing this representation, and in particular to compute the number of non-equivalent such extensions, where by $\mathfrak{G}$ equivalent to $\mathfrak{G}'$ we mean the existence of an isomorphism between them leaving $\mathfrak{H}$ and Γ fixed.

We shall think of an extension $\mathfrak{G}=\sum u_{\sigma}\mathfrak{H}$ of $\mathfrak{H}$ by Γ as the set of all symbols $u_{\sigma}A$ (σ in Γ, A in $\mathfrak{H}$) with the multiplication laws

$$Au_{\sigma} = u_{\sigma}A^{\sigma}, \qquad u_{\sigma}u_{\tau} = u_{\sigma\tau}F_{\sigma,\tau}. \tag{0.2}$$

The $F_{\sigma,\tau}$ are elements in $\mathfrak{H}$ satisfying the "associativity conditions"

$$F_{\rho,\sigma}^{\tau}F_{\rho\sigma,\tau} = F_{\rho,\sigma\tau}F_{\sigma,\tau}. \tag{0.3}$$

Such a function $F_{\sigma,\tau}$ on $\Gamma\Gamma$ to $\mathfrak{H}$ is called a factor-set(3) (f.s.) of Γ in $\mathfrak{H}$. Passage to a new set of representatives $u_{\sigma}' = u_{\sigma}C_{\sigma}$ (C_{σ} in $\mathfrak{H}$) replaces $F_{\sigma,\tau}$ by

$$F_{\sigma,\tau}' = \frac{C_{\sigma}^{\tau}C_{\tau}}{C_{\sigma\tau}}F_{\sigma,\tau}. \tag{0.4}$$

The factor-sets $F_{\sigma,\tau}$ and $F_{\sigma,\tau}'$ are called associate. Two extensions $\mathfrak{G}$ and $\mathfrak{G}'$ are equivalent if and only if the factor-sets of Γ in $\mathfrak{H}$ determined by them are associate.

If $F_{\sigma,\tau}$ and $G_{\sigma,\tau}$ are factor-sets, then so also is the product $F_{\sigma,\tau}G_{\sigma,\tau}$; for $\mathfrak{H}$ is abelian, so that (0.3) is evidently satisfied. Hence the set of all factor-sets of Γ in $\mathfrak{H}$ constitutes an abelian group $F\mathfrak{H}$. The identity element of $F\mathfrak{H}$ is the factor-set each value of which is the identity element 1 of $\mathfrak{H}$. The factor-sets associate to 1 form a subgroup $T\mathfrak{H}$ of $F\mathfrak{H}$, and the group $M\mathfrak{H}=F\mathfrak{H}/T\mathfrak{H}$ will be called the *multiplicator* of Γ in $\mathfrak{H}$. The order of $M\mathfrak{H}$ is the desired number of non-equivalent extensions of $\mathfrak{H}$ by Γ.

Let now Ω be any algebraically closed field of characteristic not dividing the order n of Γ. By a factor-set of Γ in Ω we mean a function $\omega_{\sigma,\tau}\neq 0$ on $\Gamma\Gamma$ to Ω satisfying

$$\omega_{\rho,\sigma}\omega_{\rho\sigma,\tau} = \omega_{\rho,\sigma\tau}\omega_{\sigma,\tau}. \tag{0.5}$$

---

(2) For a description of Schreier's theory of group extensions, see e.g., H. Zassenhaus, *Lehrbuch der Gruppentheorie*, pp. 89–97 (Berlin, 1937). We call $\mathfrak{G}$ an extension of $\mathfrak{H}$ by Γ if it contains $\mathfrak{H}$ as an invariant subgroup and $\mathfrak{G}/\mathfrak{H}\simeq\Gamma$.

(3) In Zassenhaus and also in MacLane and Schilling, the representative elements u_{σ} of the cosets of $\mathfrak{G}/\mathfrak{H}$ are written on the right instead of the left. This leads to the composition law $A^{\tau\rho}=(A^{\rho})^{\tau}$ instead of (0.1), and to similar slight changes in (0.3), (0.4), etc. This divergence is regrettable, but the computations are bad enough without having to keep in mind that the composition law runs backwards.

(This differs from (0.3) only in the absence of the superscript τ on the left.) $\omega_{\sigma,\tau}$ and $\omega'_{\sigma,\tau}$ are associate if

$$\omega'_{\sigma,\tau} = \frac{c_\sigma c_\tau}{c_{\sigma\tau}} \omega_{\sigma,\tau}, \qquad c_\sigma \in \Omega. \tag{0.6}$$

Schur([4]) defines the multiplicator $\mathfrak{M}$ of Γ to be the group of classes of associate factor-sets of Γ in Ω; it is independent of Ω. *We conjecture that $M\mathfrak{H}$ and $\mathfrak{M}$ are isomorphic*, and it is the purpose of the present paper to prove this conjecture in the case Γ solvable. The structure of $\mathfrak{M}$ is of course much easier to determine than that of $M\mathfrak{H}$, and has already been found by Schur for a considerable number of groups. For example, if Γ is abelian, and is represented as the direct product of cyclic groups of orders $m_1, \cdots, m_t$, then([5]) $\mathfrak{M}$ is the direct product of $\frac{1}{2}t(t-1)$ cyclic groups $\mathfrak{M}_{i,j}$ $(i<j)$, where the order of $\mathfrak{M}_{i,j}$ is the G.C.D. of m_i and m_j; the order of $\mathfrak{M}$ is thus $\prod_{i<j}(m_i, m_j)$.

If one cares to delve into the structure of $T\mathfrak{H}$, the natural thing to do is first to construct the group $V\mathfrak{H}$ of all "vectors" C_σ, which is simply the direct product of $\mathfrak{H}$ with itself n times. The mapping

$$C_\sigma \to \frac{C_\sigma^\tau C_\tau}{C_{\sigma\tau}}$$

is a homomorphism of $V\mathfrak{H}$ onto $T\mathfrak{H}$, and hence $T\mathfrak{H} \simeq V\mathfrak{H}/U\mathfrak{H}$ where $U\mathfrak{H}$ is the subgroup of $V\mathfrak{H}$ consisting of those C_σ satisfying

$$C_\sigma^\tau C_\tau = C_{\sigma\tau}.$$

We call such a function C_σ a *crossed character* (c.c.) of Γ in $\mathfrak{H}$ from the analogy with a crossed representation of Γ by matrices([6]) (i.e., by semi-linear transformations). The study of these is quite fascinating in its own right, in addition to being a necessary preliminary to the main task.

If C is a fixed element of $\mathfrak{H}$, then $C_\sigma = C^{1-\sigma}$ is a crossed character. The totality of these constitutes a subgroup $\mathfrak{H}^{1-\sigma}$ of $U\mathfrak{H}$. We make a second conjecture that the *crossed character group* $C\mathfrak{H} = U\mathfrak{H}/\mathfrak{H}^{1-\sigma}$ of Γ in $\mathfrak{H}$ is isomorphic with the group $\mathfrak{C}$ of linear characters of Γ in the field Ω (functions $\psi(\sigma)$ on Γ to Ω satisfying $\psi(\sigma\tau) = \psi(\sigma)\psi(\tau)$, $\psi(\sigma) \neq 0$), but again we are able to prove this only if Γ is solvable.

The paper is so arranged that all necessary lemmas involving explicit representation $A = H^{Z_{\sigma(\sigma)}\sigma}$ of the elements A of $\mathfrak{H}$ fall in §1. In the remainder of

([4]) I. Schur, *Über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen*, Journal für die reine und angewandte Mathematik, vol. 127 (1904), pp. 20–50.

([5]) Schur, Journal für die reine und angewandte Mathematik, vol. 132 (1907), p. 113.

([6]) A. Speiser, *Zahlentheoretische Sätze aus der Gruppentheorie*, Mathematische Zeitschrift, vol. 5 (1919), pp. 1–6; I. Schur, ibid., pp. 7–10.

the paper, only the formal properties expressed by these lemmas are used. In §2 the connection between $M\mathfrak{H}$ and $\mathfrak{M}$ is indicated by means of the notion of trace: tr $A=\sum a(\sigma)$. The main theorem is stated in this section in two parts (Theorems 1A and 2A). In §3 crossed characters are developed, and the second conjecture proved (Theorems 1B and 2B). Theorem 1A is proved in §4, Theorem 2A in §5. In the final brief section, §6, $M\mathfrak{H}_0$ and $C\mathfrak{H}_0$ are determined, where $\mathfrak{H}_0$ is the subgroup of $\mathfrak{H}$ of vanishing traces.

Actually we prove these four theorems for any solvable subgroup Γ^* of Γ, and make no restrictions on Γ. Theorem 2B does not even require that Γ^* be solvable. Theorem 1B might be called the "principal genus theorem for crossed characters of Γ^* in $\mathfrak{H}$," being in essence just an extension of Lemma 2. O. F. G. Schilling was the first to observe the importance of such a theorem in his work on the class field theory. The authors wish to acknowledge that they profited considerably from discussions with Schilling.

1. **Norms, traces, and the principal genus theorem.** Let Δ be any subgroup of Γ, the elements of which will be denoted by $\alpha, \beta, \cdots$. The Δ-norm of an element A of $\mathfrak{H}$ is defined to be

$$N_\Delta A = \prod_\alpha A^\alpha = A^{\Sigma_\alpha \alpha}$$

extended over all α in Δ. Evidently

$$N_\Delta(AB) = (N_\Delta A)(N_\Delta B).$$

The Γ-norm of every element of $\mathfrak{H}$ is 1. For if $A=H^{\Sigma a(\sigma)\sigma}$ then

$$N_\Gamma A = A^{\Sigma_\tau \tau} = H^{\Sigma_\sigma a(\sigma)\sigma \Sigma_\tau \tau} = H^{\Sigma_\tau \tau \Sigma_\sigma a(\sigma)\sigma} = 1$$

since $\sum\tau$ is in the center of the integral group ring of Γ, and $H^{\Sigma\tau}=1$.

LEMMA 1. *An element A of $\mathfrak{H}$ is invariant under a subgroup Δ of Γ if and only if it is the Δ-norm of an element B of $\mathfrak{H}$.*

Proof. If $A=N_\Delta B=B^{\Sigma\alpha}$, then since $(\sum\alpha)\beta=\sum\alpha$ for any β in Δ, we have $A^\beta=A$.

Suppose conversely that $A^\alpha=A$ for all α in Δ, and let $A=H^{\Sigma a(\sigma)\sigma}$. We then have

$$H^{\Sigma a(\sigma)\sigma} = A = A^\alpha = H^{\Sigma_\sigma a(\sigma\alpha^{-1})\sigma}.$$

Hence to each α in Δ there exists an integer $c(\alpha)$ such that

$$a(\sigma) - a(\sigma\alpha^{-1}) = c(\alpha)$$

for all σ in Γ. Taking the sum of these equations over σ we get $0=nc(\alpha)$, whence $c(\alpha)=0$ and

$$a(\sigma\alpha) = a(\sigma), \qquad \text{all } \sigma \in \Gamma, \alpha \in \Delta. \tag{1.1}$$

Let

$$\Gamma = \rho_1\Delta + \rho_2\Delta + \cdots + \rho_r\Delta.$$

Then by (1.1)

$$\sum_{\sigma \in \Gamma} a(\sigma)\sigma = \sum_{i=1}^{r} \sum_{\alpha \in \Delta} a(\rho_i)\rho_i\alpha = \sum_i a(\rho_i)\rho_i \sum_\alpha \alpha.$$

Hence if $B = H^{\Sigma_i a(\rho_i)\rho_i}$ we have

$$N_\Delta B = H^{\Sigma_i a(\rho_i)\rho_i \Sigma_\alpha \alpha} = H^{\Sigma a(\sigma)\sigma} = A,$$

which completes the proof.

Since $N_\Gamma A = 1$ for all A in $\mathfrak{H}$, it follows from this lemma that *there are no elements* $\neq 1$ *of* $\mathfrak{H}$ *invariant under* Γ.

If ρ is any element of Γ and A is in $\mathfrak{H}$, we define the ρ-norm $N_\rho A$ of A to be $N_\Delta A$, where Δ is the cyclic group $\{\rho\}$ generated by ρ. If ρ has order r, then

$$N_\rho A = A^{1+\rho+\rho^2+\cdots+\rho^{r-1}}.$$

The *trace* of an element $A = H^{\Sigma a(\sigma)\sigma}$ of $\mathfrak{H}$ is defined by

$$\operatorname{tr} A = \sum a(\sigma).$$

Since for any integer c, $A = H^{\Sigma(a(\sigma)+c)\sigma}$ we have

$$\operatorname{tr} A = \sum a(\sigma) + nc,$$

and hence tr A is actually an integer taken modulo n. The coefficients $a(\sigma)$ can be chosen so that tr A has one of the values $0, 1, 2, \cdots, n-1$. It is clear that elements of any preassigned trace exist in $\mathfrak{H}$, for example, the powers of H. Evidently

$$(1.2) \qquad \operatorname{tr} AB \equiv \operatorname{tr} A + \operatorname{tr} B, \qquad \operatorname{tr} A^\sigma \equiv \operatorname{tr} A, \qquad \text{all } A, B \in \mathfrak{H}, \sigma \in \Gamma.$$

All congruences throughout the paper are taken modulo the order n *of* Γ *and the sign* (mod n) *will be omitted.*

The next lemma we might call the "principal genus theorem" for $\mathfrak{H}$, being analogous to Hilbert's theorem on units in a cyclic field(7).

LEMMA 2. *Let* ρ *be any element of* Γ, *and let* Δ^* *be any subgroup of* Γ *invariant under* ρ. *Then an element* A *of* $\mathfrak{H}$ *has the form* $A = B^{1-\rho}$ *with* B *invariant under* Δ^* *if and only if* $N_\rho A = 1$, $\operatorname{tr} A \equiv 0$, *and* A *is invariant under* Δ^*.

Proof. Let $\Delta = \{\rho, \Delta^*\}$ be the group generated by ρ and Δ^*. Let g be the index of Δ^* in Δ, so that

$$\Delta = \Delta^* + \Delta^*\rho + \Delta^*\rho^2 + \cdots + \Delta^*\rho^{g-1}, \qquad \rho^g \in \Delta^*.$$

(7) D. Hilbert, *Gesammelte Abhandlungen*, vol. 1, Theorem 90, p. 149.

(The lemma is trivial if $\rho \in \Delta^*$; for, since $\mathfrak{P}$ has no elements $\neq 1$ of finite order, each condition is equivalent to $A = 1$.) If $A = B^{1-\rho}$ with B invariant under Δ^*, and r is the order of ρ, then

$$\operatorname{tr} A \equiv \operatorname{tr} B - \operatorname{tr} B^{\rho} \equiv 0, \qquad N_{\rho} A = B^{1-\rho^r} = B^0 = 1.$$

If α is in Δ^* then so also is $\beta = \rho\alpha\rho^{-1}$, since Δ^* is invariant under ρ. Hence we have

$$B^{\rho\alpha} = B^{\beta\rho} = B^{\rho},$$

and

$$A^{\alpha} = B^{\alpha-\rho\alpha} = B^{1-\rho} = A.$$

Assume conversely that $A = H^{\Sigma a(\sigma)\sigma}$ is invariant under Δ^*, $N_{\rho} A = 1$, and $\operatorname{tr} A \equiv 0$. As in the proof of Lemma 1, equation (1.1), the invariance of A under Δ^* is equivalent to

$$a(\sigma\alpha) = a(\sigma), \qquad \text{all } \sigma \in \Gamma, \alpha \in \Delta^*. \tag{1.3}$$

Again let r be the order of ρ. Then

$$\sum_{\sigma} a(\sigma)\sigma \sum_{i=0}^{r-1} \rho^i = \sum_{i} \sum_{\sigma} a(\sigma\rho^{-i})\sigma = \sum_{\sigma} \Big(\sum_{i} a(\sigma\rho^i) \Big) \sigma,$$

and hence $N_{\rho} A = 1$ requires that

$$\sum_{i=0}^{r-1} a(\sigma\rho^i) = c, \qquad \text{all } \sigma \in \Gamma, \tag{1.4}$$

where c is an integer independent of σ.

Let $l = n/r$, and let $\lambda_1, \cdots, \lambda_l$ be representative elements of the cosets in Γ of the cyclic group $\{\rho\}$ generated by ρ:

$$\Gamma = \lambda_1\{\rho\} + \lambda_2\{\rho\} + \cdots + \lambda_l\{\rho\}.$$

We now choose the components $a(\sigma)$ of A so that $\operatorname{tr} A$ is actually zero, not merely divisible by n. Then from (1.4).

$$0 = \operatorname{tr} A = \sum_{j=1}^{l} \sum_{i=0}^{r-1} a(\lambda_j\rho^i) = \sum_{j=1}^{l} c = lc,$$

whence $c = 0$. Using (1.3) and the fact that ρ^g is in Δ^*, (1.4) with $c = 0$ gives

$$\sum_{i=0}^{g-1} a(\sigma\rho^i) = 0, \qquad \text{all } \sigma \in \Gamma. \tag{1.5}$$

Now let

$$\Gamma = \mu_1\Delta + \mu_2\Delta + \cdots + \mu_m\Delta.$$

Every element of Γ is then uniquely expressible in the form $\mu_j\alpha\rho^i$ $(j=1, \cdots, m; \alpha\in\Delta^*; i=0, \cdots, g-1)$. We define the element $B=H^{\Sigma b(\sigma)\sigma}$ as follows:

$$b(\mu_j\alpha\rho^i) = a(\mu_j\alpha) + a(\mu_j\alpha\rho) + \cdots + a(\mu_j\alpha\rho^i).$$

(1.5) and $\rho^g\in\Delta^*$ give $b(\mu_j\alpha\rho^{-1})=0$, and hence

$$b(\mu_j\alpha\rho^i) - b(\mu_j\alpha\rho^{i-1}) = a(\mu_j\alpha\rho^i).$$

For any element $\sigma=\mu_j\alpha\rho^i$ of Γ this gives

$$b(\sigma) - b(\sigma\rho^{-1}) = a(\sigma),$$

whence $A=B^{1-\rho}$.

From (1.3),

$$a(\mu_j\alpha\rho^i) = a(\mu_j\rho^i\cdot\rho^{-i}\alpha\rho^i) = a(\mu_j\rho^i),$$

whence

$$b(\mu_j\alpha\rho^i) = b(\mu_j\rho^i).$$

Hence for any element $\sigma=\mu_j\alpha\rho^i$ of Γ and any element β of Δ^*,

$$b(\sigma\beta) = b(\mu_j\cdot\alpha\rho^i\beta\rho^{-i}\cdot\rho^i) = b(\mu_j\rho^i) = b(\sigma),$$

whence $B^\beta=B$, and the proof is complete.

LEMMA 3. *If Δ is a subgroup of index m in Γ, then there exists an element A of $\mathfrak{H}$ such that* tr $A\equiv m$ *and* $N_\Delta A=1$.

Proof. Let

$$\Gamma = \mu_1\Delta + \mu_2\Delta + \cdots + \mu_m\Delta.$$

Define $A=H^{\Sigma a(\sigma)\sigma}$ as follows:

$$a(\mu_j\alpha) = \begin{cases} 1 & \text{if } \alpha = 1, \\ 0 & \text{if } \alpha \neq 1, \end{cases} \qquad \alpha\in\Delta;\ j = 1, \cdots, m.$$

That tr $A\equiv m$ is obvious. If $\sigma=\mu_j\beta$ is any element of Γ,

$$\sum_{\alpha\in\Delta} a(\sigma\alpha^{-1}) = \sum_\alpha a(\mu_j\beta\alpha^{-1}) = \sum_\alpha a(\mu_j\alpha) = 1.$$

Hence

$$\sum_{\sigma\in\Gamma} a(\sigma)\sigma \sum_{\alpha\in\Delta} \alpha = \sum_\sigma \left(\sum_\alpha a(\sigma\alpha^{-1}) \right)\sigma = \sum_\sigma \sigma,$$

and

$$N_\Delta A = H^{\Sigma\sigma} = 1.$$

COROLLARY. *An element A of $\mathfrak{H}$ has the form $A=N_\rho B$ with* tr $B\equiv 0$ *if and only if $A^\rho=A$ and* tr $A\equiv 0$.

Proof. If $A=N_\rho B$ with tr $B\equiv 0$, then $A^\rho=A$ is evident and tr $A\equiv r$ tr $B\equiv 0$, where r is the order of ρ. Conversely, from $A^\rho=A$ follows $A=N_\rho B_1$ for some $B_1\in\mathfrak{H}$ by Lemma 1. Then r tr $B_1\equiv$ tr $A\equiv 0$. Fixing on a definite representation of B_1, this asserts that r tr $B_1=ln$ for some integer l. Applying Lemma 3 to $\Delta=\{\rho\}$, there exists C in $\mathfrak{H}$ such that tr $C\equiv n/r$, $N_\rho C=1$. Set $B=B_1C^{-l}$. Then

$$N_\rho B = (N_\rho B_1)(N_\rho C)^{-l} = N_\rho B_1 = A,$$

$$\text{tr } B \equiv \text{tr } B_1 - l \text{ tr } C \equiv \frac{ln}{r} - \frac{ln}{r} \equiv 0.$$

2. **Connection with Schur's multiplicator.** In the brief discussion of factor-sets and group extensions in the introduction, $\mathfrak{H}$ may of course be replaced by any abelian group $\mathfrak{A}$, and we may take any representation of Γ by automorphisms $A\to A^\sigma$ of $\mathfrak{A}$. The two cases with which we shall be chiefly concerned are: (1) Γ replaced by a subgroup Γ^*, without changing $\mathfrak{H}$, and (2) $\mathfrak{A}$ a cyclic group $\mathfrak{Z}$ of order n, with $z^\sigma=z$ for all $z\in\mathfrak{Z}$, $\sigma\in\Gamma$. (When Γ induces the identical automorphism in $\mathfrak{A}$ we shall speak of *central extensions* $\mathfrak{G}$ of $\mathfrak{A}$ by Γ, since $\mathfrak{A}$ is in the center of $\mathfrak{G}$.)

We shall call an f.s. $F_{\sigma,\tau}$ *normalized* if, in the extension $\mathfrak{G}$ defined by (0.2),

$$\sigma^s = 1 \text{ implies } u_\sigma^s = 1.$$

From (0.2) one readily sees that this is equivalent to

$$F_{\sigma,\sigma}F_{\sigma,\sigma^2}\cdots F_{\sigma,\sigma^{s-1}} = 1 \quad \text{if} \quad \sigma^s = 1. \tag{2.1}$$

The set of all normalized f.s.'s of Γ in $\mathfrak{A}$ constitutes a subgroup $\overline{F}\mathfrak{A}$ of $F\mathfrak{A}$, and we shall call the corresponding subgroup $\overline{M}\mathfrak{A}$ of $M\mathfrak{A}$ the *normalized multiplicator* of Γ in $\mathfrak{A}$. In general, e.g., in case (2) above, $\overline{M}\mathfrak{A}$ is a proper subgroup of $M\mathfrak{A}$, but in case (1) they coincide:

If $\mathfrak{G}^*=\sum u_\sigma\mathfrak{H}$ $(\sigma\in\Gamma^*)$ is any extension of $\mathfrak{H}$ by a subgroup Γ^* of Γ, new representatives v_σ can be found such that if $\sigma^s=1$ then $v_\sigma^s=1$.

In proving this we may evidently assume that s is the order of σ. It is clear that we can first normalize the u_σ so that $u_1=1$. Then

$$u_\sigma^s = F_{\sigma,\sigma}F_{\sigma,\sigma^2}\cdots F_{\sigma,\sigma^{s-1}} = A_\sigma \tag{2.2}$$

is an element of $\mathfrak{H}$. Since

$$A_\sigma^\sigma = u_\sigma^{-1}A_\sigma u_\sigma = u_\sigma^s = A_\sigma$$

there exists by Lemma 1 an element B_σ of $\mathfrak{H}$ such that

$$A_\sigma = N_\sigma B_\sigma = B_\sigma^{1+\sigma+\cdots+\sigma^{s-1}}. \tag{2.3}$$

If we set $u_\sigma = v_\sigma B_\sigma$ then an easy calculation shows that

$$u_\sigma^s = v_\sigma^s B_\sigma^{1+\sigma+\cdots+\sigma^{s-1}} = v_\sigma^s A_\sigma,$$

whence $v_\sigma' = 1$.

Now Schur has shown(8) that any f.s. $\omega_{\sigma,\tau}$ of Γ in the field Ω (see introduction, equation (0.5)) is associate to one such that each number $\omega_{\sigma,\tau}$ is an nth root of unity. $\omega_{\sigma,\tau}$ defines a central extension $\sum \bar{u}_\sigma \Omega^*$ of the multiplicative group Ω^* of Ω by Γ:

$$\bar{u}_\sigma \bar{u}_\tau = \bar{u}_{\sigma\tau}\omega_{\sigma,\tau}, \qquad \omega \bar{u}_\sigma = \bar{u}_\sigma \omega, \qquad \text{all } \omega \in \Omega. \tag{2.4}$$

We proceed to show that $\omega_{\sigma,\tau}$ is associate to a normalized f.s. $\omega'_{\sigma,\tau}$ which likewise has the property $\omega'^n_{\sigma,\tau} = 1$. Let

$$\omega_\sigma = \prod_\tau \omega_{\sigma,\tau}.$$

Taking the product of (0.5) over τ and using $\omega^n_{\rho,\sigma} = 1$, we get

$$\omega_{\rho\sigma} = \omega_\rho \omega_\sigma, \tag{2.5}$$

i.e., ω_σ is a linear character of Γ in Ω. If σ has order s and τ is arbitrary in Γ, then by successive multiplication of (2.4) on the left by $\bar{u}_\sigma$ we find

$$a_\sigma = \bar{u}_\sigma^s = \omega_{\sigma,\tau}\omega_{\sigma,\sigma\tau}\cdots\omega_{\sigma,\sigma^{s-1}\tau}.$$

If we think of Γ as decomposed into $\{\sigma\}$ and its cosets, we see from this that $\omega_\sigma = a_\sigma^{n/s}$. Let c_σ be any sth root of a_σ^{-1} in Ω, and let $\bar{v}_\sigma = \bar{u}_\sigma c_\sigma$. Then $\bar{v}_\sigma^s = 1$, so that the new factor-set $\omega'_{\sigma,\tau}$ (given by (0.6)) is normalized. Since $c_\sigma^n = a_\sigma^{-n/s} = \omega_\sigma^{-1}$, raising (0.6) to the nth power and using (2.5) gives $\omega'^n_{\sigma,\tau} = 1$.

Since the $\omega_{\sigma,\tau}$ lie in the cyclic group $\mathfrak{Z} = \mathfrak{Z}_n$ of nth roots of unity in Ω, the equations (2.4) also serve to define a central extension $\overline{\mathfrak{G}} = \sum \bar{u}_\sigma \mathfrak{Z}$ of $\mathfrak{Z}$ by Γ. If two normalized f.s.'s of Γ in $\mathfrak{Z}$ are associate relative to $\mathfrak{Z}$ they are plainly associate relative to Ω. But the converse is also true. For if $\bar{v}_\sigma = \bar{u}_\sigma c_\sigma$ ($c_\sigma \in \Omega$), from $\bar{u}_\sigma^n = \bar{v}_\sigma^n = 1$ follows $c_\sigma^n = 1$, i.e., $c_\sigma \in \mathfrak{Z}$. We have therefore shown that *the multiplicator $\mathfrak{M}$ of any finite group Γ of order n* (as defined by Schur) *is isomorphic with the normalized multiplicator $\overline{M}\mathfrak{Z}$ of Γ in a cyclic group $\mathfrak{Z}$ of order n.* (The normalized f.s.'s of Γ in $\mathfrak{Z}$ correspond to those central extensions $\overline{\mathfrak{G}}$ of $\mathfrak{Z}$ by Γ having the property that the order of every element of $\overline{\mathfrak{G}}$ divides n.)

Let ϵ be a primitive nth root of unity in Ω, fixed throughout the rest of

(8) In the 1904 paper, p. 26.

the paper. If $F_{\sigma,\tau}$ is any f.s. of Γ in $\mathfrak{H}$ and we set

$$f_{\sigma,\tau} = \operatorname{tr} F_{\sigma,\tau}, \tag{2.6}$$

then from (0.3) and (1.2)

$$f_{\rho,\sigma} + f_{\rho\sigma,\tau} \equiv f_{\rho,\sigma\tau} + f_{\sigma,\tau}, \tag{2.7}$$

and hence

$$\omega_{\sigma,\tau} = e^{f_{\sigma,\tau}} \tag{2.8}$$

is an f.s. of Γ in $\mathfrak{Z}$. By taking traces of (0.4) it is clear that associate f.s.'s of Γ in $\mathfrak{H}$ give rise to associate f.s.'s of Γ in $\mathfrak{Z}$:

$$f'_{\sigma,\tau} \equiv c_\sigma + c_\tau - c_{\sigma\tau} + f_{\sigma,\tau}, \qquad c_\sigma = \operatorname{tr} C_\sigma. \tag{2.9}$$

Again by (1.2) a product FF' maps into the sum $f+f'$ and thence into the product $\omega\omega'$. Finally, by taking the trace of (2.1),

$$f_{\sigma,\sigma} + f_{\sigma,\sigma^2} + \cdots + f_{\sigma,\sigma^{s-1}} \equiv 0, \qquad \sigma^s = 1, \tag{2.10}$$

and using (2.8) it is seen that a normalized F maps into a normalized ω.

Since every F has been shown to be associate to a normalized one, the mapping $F \to f \to \omega$ given by (2.6) and (2.8) is a homomorphism of $M\mathfrak{H}$ onto a subgroup of $\overline{M}\mathfrak{Z}$. As mentioned in the introduction, our conjecture is that this is actually an isomorphism between $M\mathfrak{H}$ and $\overline{M}\mathfrak{Z}$ ($\simeq \mathfrak{M}$), and we shall prove it for solvable Γ. We must show (1) that distinct elements of $M\mathfrak{H}$ correspond to distinct elements of $\overline{M}\mathfrak{Z}$, and (2) that $\overline{M}\mathfrak{Z}$ is covered by the mapping. We accordingly split the main theorem into two parts. All the foregoing holds as well for any subgroup Γ^* of Γ, and the inductive nature of the proof makes it necessary to consider this more general situation. We are of course rewarded with a correspondingly more general result; only the subgroup Γ^* need be solvable, Γ being perfectly arbitrary.

To avoid circumlocution, we shall identify $\mathfrak{Z}$ with the additive group of integers mod n. An f.s. of Γ in $\mathfrak{Z}$ is then a set of integers $f_{\sigma,\tau}$ satisfying the congruence (2.7). Associate f.s.'s are connected as in (2.9). (2.10) is the normalization condition.

THEOREM 1A. *If Γ^* is a solvable subgroup of Γ, and if $F_{\sigma,\tau}$ is a factor-set of Γ^* in $\mathfrak{H}$ such that the factor-set* tr $F_{\sigma,\tau}$ *of Γ^* in $\mathfrak{Z}$ is associate to* 1, *then $F_{\sigma,\tau}$ is itself associate to* 1.

THEOREM 2A. *If Γ^* is a solvable subgroup of Γ, and if $f_{\sigma,\tau}$ is a normalized factor-set of Γ^* in $\mathfrak{Z}$, then there exists a normalized factor-set $F_{\sigma,\tau}$ of Γ^* in $\mathfrak{H}$ such that* tr $F_{\sigma,\tau} \equiv f_{\sigma,\tau}$.

Theorem 1A will be proved in §4, Theorem 2A in §5. It is first necessary to develop the theory of crossed characters.

3. **Crossed characters.** Let Γ^* be a subgroup of order m of Γ which we regard as fixed throughout this section. There will be no occasion to consider elements of Γ outside of Γ^*.

As mentioned in the introduction, a function A_σ on Γ^* to $\mathfrak{H}$ will be called a crossed character (c.c.) of Γ^* in $\mathfrak{H}$ if

$$A_\sigma^\tau A_\tau = A_{\sigma\tau}. \tag{3.1}$$

If B_σ is another c.c. of Γ^* in $\mathfrak{H}$, so is $A_\sigma B_\sigma$, and likewise A_σ^{-1}. Hence they form an abelian group $U\mathfrak{H}$.

If A is any element of $\mathfrak{H}$ and we set $A_\sigma = A^{1-\sigma}$ for all σ in Γ^*, then A_σ is a c.c. For

$$A_\sigma^\tau A_\tau = A^{(1-\sigma)\tau+(1-\tau)} = A^{1-\sigma\tau} = A_{\sigma\tau}.$$

$A^{1-\sigma}$ will be called a *unit* c.c. The unit c.c.'s of Γ^* in $\mathfrak{H}$ constitute a subgroup $\mathfrak{H}^{1-\sigma}$ of $U\mathfrak{H}$. (Incidentally(9), by taking the product of (3.1) over σ we find that the mth power of any c.c. is a unit c.c.)

If $a_\sigma = \operatorname{tr} A_\sigma$ then from (3.1)

$$a_\sigma + a_\tau \equiv a_{\sigma\tau} \tag{3.2}$$

and hence

$$\psi_A(\sigma) = \epsilon^{a_\sigma}$$

is a linear character of Γ^* in the field Ω. Evidently

$$\psi_{AB}(\sigma) = \psi_A(\sigma)\psi_B(\sigma).$$

Moreover, if $A_\sigma = A^{1-\sigma}$ then $\operatorname{tr} A_\sigma \equiv 0$ and $\psi_A(\sigma) = 1$. Hence the mapping $A_\sigma \to \psi_A(\sigma)$ is a homomorphism of the crossed character group $C\mathfrak{H} = U\mathfrak{H}/\mathfrak{H}^{1-\sigma}$ of Γ^* in $\mathfrak{H}$ onto a subgroup of the group $\mathfrak{C}$ of linear characters of Γ^* in Ω. ($\mathfrak{C}$ is of course identical with the character group $C\mathfrak{Z}$ of Γ^* in $\mathfrak{Z}$.) We conjecture that this mapping is an isomorphism between $C\mathfrak{H}$ and $\mathfrak{C}$, and shall prove it for the case Γ^* solvable. Again we split the proof into its two parts. Observe here that *the second part* (Theorem 2B) *does not require the solvability of* Γ^*.

THEOREM 1B. *If Γ^* is a solvable subgroup of Γ, and if A_σ is a crossed character of Γ^* in $\mathfrak{H}$ such that* $\operatorname{tr} A_\sigma \equiv 0$, *then A_σ is a unit crossed character.*

THEOREM 2B. *If Γ^* is any subgroup of Γ, and $\psi(\sigma)$ is a linear character of Γ^* in Ω, then there exists a crossed character A_σ of Γ^* in $\mathfrak{H}$ such that $\epsilon^{\operatorname{tr} A_\sigma} = \psi(\sigma)$.*

If we identify $\mathfrak{Z}$ with the additive group of integers mod n, Theorem 2B is more conveniently expressed as follows. If Γ^* is any subgroup of Γ, and a_σ is a

(9) Cf. Zassenhaus, loc. cit., p. 126. Thus every c.c. of a finite group of order m in an abelian group of finite exponent prime to m is a unit c.c.

linear character of Γ^* in $\mathfrak{Z}$, i.e., the a_σ are integers satisfying (3.2), then there exists a c.c. A_σ of Γ^* in $\mathfrak{H}$ such that $\operatorname{tr} A_\sigma \equiv a_\sigma$.

Proof of 1B. Let A_σ be a c.c. of Γ^* in $\mathfrak{H}$. Putting $\sigma=\tau=1$ in (3.1) we find $A_1=1$. From (3.1) one readily computes

$$A_{\sigma^i} = A_\sigma^{1+\sigma+\sigma^2+\cdots+\sigma^{i-1}} \tag{3.3}$$

for any σ in Γ^* and any positive integer i. If s is the order of σ, then $i=s$ yields

$$N_\sigma A_\sigma = 1. \tag{3.4}$$

We prove the theorem by showing that A_σ can be reduced to 1 by successive multiplication by unit c.c.'s $B^{1-\sigma}$. The result will be established by an obvious induction on a composition series of Γ^* when we have proved the following statement.

Let $\Delta^* \subset \Delta$ be subgroups of Γ^* such that Δ^* is invariant in Δ, and Δ/Δ^* is cyclic. If $A_\alpha=1$ for all α in Δ^*, and $\operatorname{tr} A_\sigma \equiv 0$ for all σ in Δ, then there exists B in $\mathfrak{H}$ such that $A_\sigma = B^{1-\sigma}$ for all σ in Δ.

Let

$$\Delta = \Delta^* + \Delta^*\rho + \Delta^*\rho^2 + \cdots + \Delta^*\rho^{\varrho-1}, \qquad \rho^\varrho \in \Delta^*.$$

From (3.1) and the condition $A_\alpha=1$ we have, for every α in Δ^*,

$$A_{\alpha\rho} = A_\rho, \qquad A_{\rho\alpha} = A_\rho^\alpha,$$

and, since $\rho\alpha\rho^{-1}$ is in Δ^*,

$$A_\rho^\alpha = A_{\rho\alpha} = A_{\rho\alpha\rho^{-1}\cdot\rho} = A_\rho.$$

Hence A_ρ is invariant under Δ^*, $N_\rho A_\rho = 1$ by (3.4), and $\operatorname{tr} A_\rho \equiv 0$ by hypothesis. By Lemma 2 of §1 there exists an element $B=B_\rho$ of $\mathfrak{H}$ (ρ is a fixed element!) such that

$$A_\rho = B^{1-\rho}, \qquad B^\alpha = B, \qquad \text{all } \alpha \in \Delta^*.$$

By (3.3),

$$A_{\rho^i} = B^{1-\rho^i},$$

and hence for any $\sigma=\alpha\rho^i$ in Δ,

$$A_\sigma = A_{\alpha\rho^i} = A_{\rho^i} = B^{1-\rho^i} = B^{1-\alpha\rho^i} = B^{1-\sigma}.$$

Proof of 2B. Let Γ' be the commutator subgroup of Γ^*, and let $\rho_1, \cdots, \rho_t$ be representative elements of basic classes of the abelian group Γ^*/Γ' (expressed as a direct product of t cyclic groups). Let r_j be the order of $\rho_j \bmod \Gamma'$. Every linear character of Γ^* is a product of the basic characters ψ_j ($j=1, \cdots, t$) defined in the obvious way from

$$\psi_j(\rho_j) = \epsilon^{n/r_j}, \qquad \psi_j(\rho_k) = 1 \quad \text{for} \quad k \neq j.$$

(Any linear character of Γ^* must have the value 1 over Γ'.) We need only show that each basic ψ_j can be realized by a c.c. $A_\sigma^{(j)}$ of Γ^* in $\mathfrak{H}$, for then any linear character ψ can be realized by taking the corresponding product of the $A_\sigma^{(j)}$.

Without loss in generality we may take $j=1$, and for simplicity we write ρ for ρ_1, r for r_1, ψ for ψ_1. $\Delta=\{\Gamma', \rho_2, \cdots, \rho_t\}$ is an invariant subgroup of Γ^*,

$$\Gamma^* = \Delta + \Delta\rho + \Delta\rho^2 + \cdots + \Delta\rho^{r-1}, \qquad \rho^r \in \Gamma' \subseteq \Delta,$$

and for each $\sigma=\alpha\rho^i$ of Γ^* ($\alpha\in\Delta$; $i=0, 1, \cdots, r-1$) we have

$$\psi(\alpha\rho^i) = \epsilon^{in/r}. \tag{3.5}$$

By Lemma 3 of §1 there exists an element B of $\mathfrak{H}$ such that

$$\operatorname{tr} B \equiv \frac{n}{m}, \qquad N_{\Gamma^*}B = 1,$$

since n/m is the index of Γ^* in Γ. Let

$$A = N_\Delta B.$$

Since m/r is the order of Δ,

$$\operatorname{tr} A \equiv \frac{m}{r} \operatorname{tr} B \equiv \frac{m}{r}\cdot\frac{n}{m} \equiv \frac{n}{r}. \tag{3.6}$$

Also, since

$$\sum_{\alpha\,\varepsilon\,\Delta} \alpha \sum_{i=0}^{r-1} \rho^i = \sum_{\sigma\,\varepsilon\,\Gamma^*} \sigma,$$

we have

$$A^{1+\rho+\cdots+\rho^{r-1}} = N_{\Gamma^*}B = 1. \tag{3.7}$$

For every $\sigma=\alpha\rho^i$ in Γ^* we now define

$$A_{\alpha\rho^i} = A^{1+\rho+\cdots+\rho^{i-1}}. \tag{3.8}$$

Since A is a Δ-norm, it is invariant under the element ρ^r of Δ. From this and (3.7) it follows that (3.8) is valid for all integers $i\geqq 0$. If β is in Δ so also are $\beta_j=\rho^j\beta\rho^{-j}$ ($j=0, 1, \cdots, i-1$), and, since A is invariant under Δ,

$$A_{\alpha\rho^i}^{\beta} = A^{\beta_0+\beta_1\rho+\cdots+\beta_{i-1}\rho^{i-1}} = A_{\alpha\rho^i}.$$

For any two elements $\sigma=\alpha\rho^i$, $\tau=\beta\rho^j$ of Γ^* we therefore have

$$A_{\sigma\tau} = A_{\alpha\rho^i\beta\rho^{-i}\cdot\rho^{i+j}} = A_{\rho^{i+j}},$$

$$A_\sigma^\tau A_\tau = A_{\alpha\rho^i\beta\rho^j}^{\rho^j} A_{\beta\rho^j} = A^{(1+\rho+\cdots+\rho^{i-1})\rho^j+(1+\rho\cdots+\rho^{j-1})} = A_{\rho^{i+j}}.$$

Hence A_σ defined by (3.8) is a c.c. of Γ^* in $\mathfrak{H}$. By (3.6) for $\sigma=\alpha\rho^i$,

$$\operatorname{tr} A_\sigma \equiv i \operatorname{tr} A \equiv \frac{in}{r},$$

whence by (3.5),

$$\epsilon^{\operatorname{tr} A_\sigma} = \psi(\sigma).$$

Hence A_σ realizes the basic character $\psi(\sigma)$.

4. **Proof of Theorem** 1A. By hypothesis, Γ^* is a fixed solvable subgroup of Γ and $F_{\sigma,\tau}$ is an f.s. of Γ^* in $\mathfrak{H}$ such that

$$\operatorname{tr} F_{\sigma,\tau} \equiv c_\sigma + c_\tau - c_{\sigma\tau}, \qquad \text{all } \sigma, \tau \in \Gamma^*, \tag{4.1}$$

where the c_σ are integers. By an obvious induction on a composition series of Γ^*, the theorem will be established when we have proved the following statement.

Let $\Delta^* \subset \Delta$ be subgroups of Γ^* such that Δ^* is invariant in Δ and Δ/Δ^* is cyclic. If the f.s. $F_{\alpha,\beta}$ of Δ^* in $\mathfrak{H}$ is associate to 1 ($\alpha, \beta \in \Delta^*$), then the f.s. $F_{\sigma,\tau}$ of Δ in $\mathfrak{H}$ is likewise associate to 1 ($\sigma, \tau \in \Delta$).

$F_{\sigma,\tau}$ determines an extension $\mathfrak{G}=\sum u_\sigma \mathfrak{H}$ ($\sigma \in \Delta$) of $\mathfrak{H}$ by Δ:

$$u_\sigma u_\tau = u_{\sigma\tau} F_{\sigma,\tau}, \qquad A u_\sigma = u_\sigma A^\sigma.$$

By hypothesis for induction we may assume the u_α ($\alpha \in \Delta^*$) so chosen that

$$u_\alpha u_\beta = u_{\alpha\beta}, \qquad \alpha, \beta \in \Delta^*. \tag{4.2}$$

(It is hardly necessary to point out that an f.s. $F'_{\alpha,\beta}$ associate to $F_{\alpha,\beta}$ within a subgroup of Γ^* can be extended to an f.s. $F'_{\sigma,\tau}$ of Γ^* associate to $F_{\sigma,\tau}$, and that (4.1) remains valid with new integers c'_σ.)

Let

$$\Delta = \Delta^* + \rho\Delta^* + \cdots + \rho^{g-1}\Delta^*, \qquad \rho^g \in \Delta^*.$$

Then

$$u_\rho^{-1} u_\alpha u_\rho = u_{\rho^{-1}\alpha\rho} A_{\rho^{-1}\alpha\rho}, \qquad \text{all } \alpha \in \Delta^*, \tag{4.3a}$$

$$u_\rho^g = u_{\rho^g} B, \tag{4.3b}$$

where the elements $A_{\rho^{-1}\alpha\rho}$ and B are in $\mathfrak{H}$. The proof consists in showing that these elements can be reduced to 1 by successive normalizations of the u_σ. For if, along with (4.2), we have the equations

$$u_\rho^{-1} u_\alpha u_\rho = u_{\rho^{-1}\alpha\rho}, \qquad u_\rho^g = u_{\rho^g}$$

and define for any $\sigma=\rho^i\alpha$ of Δ ($\alpha \in \Delta^*$; $i=0, 1, \cdots, g-1$)

$$u_{\rho^i\alpha} = u_\rho^i u_\alpha,$$

it is readily seen this last holds for all integers $i \geqq 0$, and for any $\sigma = \rho^i\alpha$, $\tau = \rho^j\beta$ of Δ,

$$u_\sigma u_\tau = u_\rho^{i+j} u_\rho^{-j} u_\alpha u_\rho^j u_\beta = u_\rho^{i+j} u_{\rho^{-j}\alpha\rho^j} u_\beta = u_{\rho^{i+j}\cdot\rho^{-j}\alpha\rho^j\beta} = u_{\sigma\tau}.$$

This means, of course, that the original $F_{\sigma,\tau}$ is associate to 1.

(4.3a) may also be written

$$u_\rho^{-1} u_{\rho\alpha\rho^{-1}} u_\rho = u_\alpha A_\alpha.$$

Hence for $\alpha, \beta \in \Delta^*$, using (4.2),

$$u_{\alpha\beta} A_{\alpha\beta} = u_\rho^{-1} u_{\rho\alpha\rho^{-1}} u_{\rho\beta\rho^{-1}} u_\rho = u_\alpha A_\alpha u_\beta A_\beta = u_{\alpha\beta} A_\alpha^\beta A_\beta.$$

A_α is thus a crossed character of Δ^* in $\mathfrak{H}$. We proceed to show that the u_α can be normalized so that $\operatorname{tr} A_\alpha \equiv 0$. From

$$u_\alpha u_\rho = u_{\alpha\rho} F_{\alpha,\rho}, \qquad u_\rho u_{\rho^{-1}\alpha\rho} = u_{\alpha\rho} F_{\rho,\rho^{-1}\alpha\rho}$$

and (4.3a) we find that

$$A_{\rho^{-1}\alpha\rho} = \frac{F_{\alpha,\rho}}{F_{\rho,\rho^{-1}\alpha\rho}}.$$

Hence by (4.1)

$$\operatorname{tr} A_{\rho^{-1}\alpha\rho} \equiv c_\alpha - c_{\rho^{-1}\alpha\rho}.$$

Again from (4.1) and $F_{\alpha,\beta} = 1$ we see that c_α is a linear character of Δ^* in $\mathfrak{Z}$, and hence by Theorem 2B there exists a c.c. C_α of Δ^* in $\mathfrak{H}$ with $\operatorname{tr} C_\alpha \equiv c_\alpha$. If we now set $u_\alpha = v_\alpha C_\alpha$, (4.2) remains valid for the v_α and

$$u_\rho^{-1} v_\alpha u_\rho = v_{\rho^{-1}\alpha\rho} A'_{\rho^{-1}\alpha\rho}$$

where

$$A'_{\rho^{-1}\alpha\rho} = C_{\rho^{-1}\alpha\rho} A_{\rho^{-1}\alpha\rho} C_\alpha^{-\rho}.$$

Hence

$$\operatorname{tr} A'_{\rho^{-1}\alpha\rho} \equiv c_{\rho^{-1}\alpha\rho} + \operatorname{tr} A_{\rho^{-1}\alpha\rho} - c_\alpha \equiv 0.$$

We may therefore assume that $\operatorname{tr} A_\alpha \equiv 0$ in (4.3a).

But now A_α is a c.c. of Δ^* in $\mathfrak{H}$ with zero trace, and by Theorem 1B there exists A in $\mathfrak{H}$ such that $A_\alpha = A^{1-\alpha}$. If we now set $u_\rho = v_\rho A$ we find

$$v_\rho^{-1} u_\alpha v_\rho = A u_{\rho^{-1}\alpha\rho} A_{\rho^{-1}\alpha\rho} A^{-1} = u_{\rho^{-1}\alpha\rho} A_{\rho^{-1}\alpha\rho} A^{\rho^{-1}\alpha\rho - 1} = u_{\rho^{-1}\alpha\rho}.$$

Hence we can assume that the u_σ have been normalized so that, along with (4.2),

$$u_\rho^{-1} u_\alpha u_\rho = u_{\rho^{-1}\alpha\rho}, \qquad \alpha \in \Delta^*, \tag{4.4a}$$

(4.4b) $$u_\rho^\varrho = u_{\rho^\varrho} B.$$

For $\alpha = \rho^\varrho$, (4.4a) shows that u_ρ commutes with u_{ρ^ϱ}. Hence by (4.4b) u_ρ must commute with B, so that $B^\rho = B$. By successive application of (4.4a),

$$u_\rho^{-\varrho} u_\alpha u_\rho^\varrho = u_{\rho^{-\varrho}\alpha\rho^\varrho}.$$

By (4.4b),

$$u_\alpha u_{\rho^\varrho} B = u_{\rho^\varrho} B u_{\rho^{-\varrho}\alpha\rho^\varrho}, \qquad u_{\alpha\rho^\varrho} B = u_{\alpha\rho^\varrho} B^{\rho^{-\varrho}\alpha\rho^\varrho}.$$

Replacing α by $\rho^\varrho\alpha\rho^{-\varrho}$, this gives $B^\alpha = B$ (all $\alpha \in \Delta^*$). Since B is invariant under ρ as well as Δ^*, it is invariant under Δ, and hence by Lemma 1 of §1 there exists C in $\mathfrak{H}$ such that $B = N_\Delta C$. Since

$$\sum_{\sigma \in \Delta} \sigma = \sum_{\alpha \in \Delta^*} \alpha \sum_{i=0}^{\varrho-1} \rho^i$$

it follows that if we set $D = N_{\Delta^*} C$ then

$$B = D^{1+\rho+\cdots+\rho^{\varrho-1}}.$$

Hence if we set $u_\rho = v_\rho D$, we have

$$u_\rho^\varrho = v_\rho^\varrho B, \qquad v_\rho^\varrho = u_{\rho^\varrho},$$

and (4.4a) remains valid:

$$v_\rho^{-1} u_\alpha v_\rho = D u_{\rho^{-1}\alpha\rho} D^{-1} = u_{\rho^{-1}\alpha\rho}$$

since D (being a Δ^*-norm) is invariant under Δ^*. We have thus achieved the desired normalization.

5. **Proof of Theorem** 2A. Let us make first a preliminary observation. As stated in §2, Theorem 2A apparently gives more precise information than is really required. Given the normalized f.s. $f_{\sigma,\tau}$ of Γ^* in $\mathfrak{Z}$, all we require is an f.s. $F'_{\sigma,\tau}$ of Γ^* in $\mathfrak{H}$ such that tr $F'_{\sigma,\tau}$ is associate to $f_{\sigma,\tau}$:

$$\operatorname{tr} F'_{\sigma,\tau} \equiv f_{\sigma,\tau} + c_\sigma + c_\tau - c_{\sigma\tau}.$$

But if such an $F'_{\sigma,\tau}$ is given, there exists a normalized f.s. $F_{\sigma,\tau}$ associate to it such that tr $F_{\sigma,\tau} = f_{\sigma,\tau}$. We first pick any elements C_σ in $\mathfrak{H}$ such that tr $C_\sigma = c_\sigma$. Then the associate f.s. $F_{\sigma,\tau}$ given by (0.4) has the property that tr $F_{\sigma,\tau} = f_{\sigma,\tau}$. Now taking the trace of (2.2) and using the hypothesis (2.10) that $f_{\sigma,\tau}$ is normalized, we see that the elements A_σ defined by (2.2) have zero trace. By the corollary to Lemma 3 of §1, the elements B_σ in (2.3) can be chosen so as to have zero trace. Hence on setting $u_\sigma = v_\sigma B_\sigma$ the resulting normalized f.s. has the same trace as $F_{\sigma,\tau}$.

Theorem 2A will be established by an obvious induction on a composition series of Γ^* when we have proved the following statement.

Let $\Delta^* \subset \Delta$ be subgroups of Γ^* such that Δ^* is invariant in Δ and Δ/Δ^* is cyclic. If $F_{\alpha,\beta}$ is a normalized f.s. of Δ^* in $\mathfrak{H}$ such that tr $F_{\alpha,\beta} \equiv f_{\alpha,\beta}$, then it is possible to construct a normalized f.s. $F_{\sigma,\tau}$ of Δ in $\mathfrak{H}$ such that tr $F_{\sigma,\tau} \equiv f_{\sigma,\tau}$.

By virtue of the preliminary observation (applied to Δ) *we need only construct an f.s. $F_{\sigma,\tau}$ of Δ in $\mathfrak{H}$ such that* tr $F_{\sigma,\tau}$ *is associate to $f_{\sigma,\tau}$.*

Let

$$\Delta = \Delta^* + \rho\Delta^* + \cdots + \rho^{g-1}\Delta^*, \qquad \rho^g \in \Delta^*.$$

Let $\mathfrak{A}$ be any abelian group, which we later specialize to be $\mathfrak{H}$ or $\mathfrak{Z}$, and let $A \to A^\sigma$ be a representation of Δ by automorphisms of $\mathfrak{A}$: $(A^\sigma)^\tau = A^{\sigma\tau}$. Let $\mathfrak{G}^* = \sum u_\alpha \mathfrak{A}$ $(\alpha \in \Delta^*)$ be a given extension of $\mathfrak{A}$ by Δ^*:

$$u_\alpha u_\beta = u_{\alpha\beta} F_{\alpha,\beta}, \qquad A u_\alpha = u_\alpha A^\alpha.$$

Let u_ρ be a new symbol. Then (Zassenhaus, loc. cit., p. 94) the equations

$$u_\rho^{-1} u_\alpha u_\rho = u_{\rho^{-1}\alpha\rho} A_{\rho^{-1}\alpha\rho}, \qquad \text{all } \alpha \in \Delta^*, \tag{5.1a}$$

$$u_\rho^{-1} A u_\rho = A^\rho, \qquad \text{all } A \in \mathfrak{A}, \tag{5.1b}$$

$$u_\rho^g = u_{\rho^g} B \tag{5.1c}$$

with $A_{\rho^{-1}\alpha\rho}$ and B in $\mathfrak{A}$, will define a cyclic extension

$$\mathfrak{G} = \mathfrak{G}^* + u_\rho \mathfrak{G}^* + u_\rho^2 \mathfrak{G}^* + \cdots + u_\rho^{g-1} \mathfrak{G}^*$$

of $\mathfrak{G}^*$ if and only if the following three conditions are satisfied.

I. The mapping (combining (5.1a) and (5.1b))

$$u_\alpha A \to \mathrm{P}(u_\alpha A) = u_{\rho^{-1}\alpha\rho} A_{\rho^{-1}\alpha\rho} A^\rho$$

is an automorphism of $\mathfrak{G}^*$ $(\alpha \in \Delta^*, A \in \mathfrak{A})$.

II. The gth power P^g of P is the inner automorphism defined by $u_{\rho^g} B$:

$$\mathrm{P}^g(u_\alpha A) = (u_{\rho^g} B)^{-1} u_\alpha A (u_{\rho^g} B).$$

III. $u_{\rho^g} B$ is invariant under P:

$$\mathrm{P}(u_{\rho^g} B) = u_{\rho^g} B.$$

If these conditions are satisfied the group $\mathfrak{G}$ will be an extension of $\mathfrak{A}$ by Δ. For by (5.1b) $\mathfrak{A}$ is invariant in $\mathfrak{G}$, every element of $\mathfrak{G}$ is uniquely expressible in the form

$$u_\rho^i u_\alpha A, \qquad i = 0, 1, \cdots, g-1; \alpha \in \Delta^*; A \in \mathfrak{A},$$

and if for each $\sigma = \rho^i\alpha$ in Δ we define u_σ by $u_{\rho^i\alpha} = u_\rho^i u_\alpha$, the product $u_\sigma u_\tau$ will differ from $u_{\sigma\tau}$ by a factor in $\mathfrak{A}$.

It can be shown([10]) by straightforward computation that I, II, III are equivalent respectively to the following conditions on the elements A_α, B of $\mathfrak{A}$.

(Ia) $$F^{\rho}_{\rho\alpha\rho-1,\rho\beta\rho-1} = \frac{A^{\beta}_{\alpha}A_{\beta}}{A_{\alpha\beta}}F_{\alpha,\beta}, \qquad \alpha, \beta \in \Delta^*.$$

(IIa) $$B^{1-\alpha} = \frac{F_{\rho\varrho,\alpha}}{F_{\rho\varrho\alpha\rho-\varrho,\rho\varrho}}A_{\alpha}A^{\rho}_{\rho\alpha\rho-1}A^{\rho^2}_{\rho^2\alpha\rho-2}\cdots A^{\rho\varrho-1}_{\rho\varrho-1\alpha\rho-\varrho+1}, \qquad \alpha \in \Delta^*,$$

(IIIa) $$B^{1-\rho} = A_{\rho\varrho}.$$

By hypothesis, $\omega_{\sigma,\tau} = \epsilon^{f_{\sigma,\tau}}$ is, with σ and τ ranging over Δ, an f.s. of Δ in $\mathfrak{Z}$. Let $\overline{\mathfrak{G}} = \sum \hat{u}_\sigma \mathfrak{Z}$ be the corresponding central extension of $\mathfrak{Z}$ by Δ:

$$\hat{u}_\sigma\hat{u}_\tau = \hat{u}_{\sigma\tau}\omega_{\sigma,\tau}, \qquad z\hat{u}_\sigma = \hat{u}_\sigma z, \qquad \text{all } z \in \mathfrak{Z}.$$

The equations (5.1) for $\mathfrak{A} = \mathfrak{Z}$ have the form

(5.2a) $$\hat{u}_\rho^{-1}\hat{u}_\alpha\hat{u}_\rho = \hat{u}_{\rho-1\alpha\rho}\epsilon^{a_{\rho-1\alpha\rho}}, \qquad \text{all } \alpha \in \Delta^*,$$

(5.2b) $$\hat{u}_\rho^{-1} z\hat{u}_\rho = z, \qquad \text{all } z \in \mathfrak{Z},$$

(5.2c) $$\hat{u}_\rho^{\varrho} = \hat{u}_{\rho\varrho}\epsilon^{b},$$

with integers a_α and b. Since these equations do in fact define an extension $\overline{\mathfrak{G}}$ of $\overline{\mathfrak{G}}^* = \sum \hat{u}_\alpha \mathfrak{Z}$ ($\alpha \in \Delta^*$), the conditions (Ia), (IIa), (IIIa) must hold. Translated into congruences involving the exponents $f_{\alpha,\beta}$, a_α, b of ϵ they are respectively as follows:

(Ib) $$f_{\rho\alpha\rho-1,\rho\beta\rho-1} \equiv a_\alpha + a_\beta - a_{\alpha\beta} + f_{\alpha,\beta}, \qquad \alpha, \beta \in \Delta^*,$$

(IIb) $$f_{\rho\varrho\alpha\rho-\varrho,\rho\varrho} - f_{\rho\varrho,\alpha} \equiv a_\alpha + a_{\rho\alpha\rho-1} + a_{\rho^2\alpha\rho-2} + \cdots + a_{\rho\varrho-1\alpha\rho-\varrho+1}, \qquad \alpha \in \Delta^*,$$

(IIIb) $$a_{\rho\varrho} \equiv 0.$$

The integer b does not enter into these, but from (5.2c) and the assumption that $f_{\sigma,\tau}$ is normalized we have

$$1 = \hat{u}_\rho^{n} = (\hat{u}_{\rho\varrho})^{n/\varrho}\epsilon^{nb/\varrho} = \epsilon^{nb/\varrho}.$$

Therefore

(IVb) the integer b is divisible by the index g of Δ^* in Δ.

By hypothesis for induction, we have already constructed a normalized f.s. $F_{\alpha,\beta}$ of Δ^* in $\mathfrak{H}$ such that tr $F_{\alpha,\beta} \equiv f_{\alpha,\beta}$, and corresponding thereto we have an extension $\mathfrak{G}^* = \sum u_\alpha \mathfrak{H}$ of $\mathfrak{H}$ by Δ^*. The whole proof will consist in showing that we can find elements A_α, B of $\mathfrak{H}$ satisfying (Ia), (IIa), (IIIa) (with $\mathfrak{A} = \mathfrak{H}$) and such that

([16]) Due to its rather unwieldly nature, much of the computation in this section will be omitted.

(5.3) $$\operatorname{tr} A_\alpha \equiv a_\alpha, \qquad \operatorname{tr} B \equiv b.$$

For, if such elements exist, the equations (5.1) define an extension $\mathfrak{G}=\sum u_\sigma \mathfrak{H}$ of $\mathfrak{H}$ by Δ. We then define, for each $\sigma=\rho^i\alpha$ of Δ,

$$u_{\rho^i\alpha} = u_\rho^i u_\alpha, \qquad i = 0, 1, \cdots, g-1,$$

and at the same time make a new choice of representative elements

$$\hat{u}_{\rho^i\alpha} = \hat{u}_\rho^i \hat{u}_\alpha$$

in the extension $\overline{\mathfrak{G}}$ of $\mathfrak{Z}$ by Δ. The latter determine an f.s. $f'_{\sigma,\tau}$ of Δ in $\mathfrak{Z}$ associate to $f_{\sigma,\tau}$ and coinciding with it within Δ^*. The former determine an f.s. $F_{\sigma,\tau}$ of Δ in $\mathfrak{H}$ coinciding with $F_{\alpha,\beta}$ within Δ^*. But the f.s. $\omega'_{\sigma,\tau}=\epsilon f'_{\sigma,\tau}$ of Δ in $\mathfrak{Z}$ is computed from the elements $\epsilon^a\alpha$, ϵ^b in formally the same way that $F_{\sigma,\tau}$ is computed from A_α, B. Hence the trace relations (5.3) will carry over into $\operatorname{tr} F_{\sigma,\tau}\equiv f'_{\sigma,\tau}$ for all σ and τ in Δ, which is the desired result.

$F_{\alpha,\beta}$ being an f.s. of Δ^* in $\mathfrak{H}$, that is,

(5.4) $$F^\gamma_{\alpha,\beta}F_{\alpha\beta,\gamma} = F_{\alpha,\beta\gamma}F_{\beta,\gamma}, \qquad \text{all } \alpha, \beta, \gamma \in \Delta^*,$$

one readily verifies by direct substitution in (5.4) that $F^\rho_{\rho\alpha\rho^{-1},\rho\beta\rho^{-1}}$ (ρ is a fixed element!) is also an f.s. of Δ^* in $\mathfrak{H}$. (Were the group $\mathfrak{G}$ already constructed, it would belong to the representatives $v_\alpha=u_\rho^{-1}u_{\rho\alpha\rho^{-1}}u_\rho$.) (Ib) asserts that the f.s.

$$\frac{F^\rho_{\rho\alpha\rho^{-1},\rho\beta\rho^{-1}}}{F_{\alpha,\beta}}$$

has trace $a_\alpha+a_\beta-a_{\alpha\beta}$. It follows from Theorem 1A (applied to Δ) that this f.s. is associate to 1. Hence there exist elements C_α in $\mathfrak{H}$ such that

$$F^\rho_{\rho\alpha\rho^{-1},\rho\beta\rho^{-1}} = \frac{C^\beta_\alpha C_\beta}{C_{\alpha\beta}} F_{\alpha,\beta}, \qquad \alpha, \beta \in \Delta^*.$$

Let $c_\alpha=\operatorname{tr} C_\alpha$. Then, taking traces,

$$f_{\rho\alpha\rho^{-1},\rho\beta\rho^{-1}} \equiv c_\alpha + c_\beta - c_{\alpha\beta} + f_{\alpha,\beta}.$$

Combining this with (Ib), we see that $d_\alpha=a_\alpha-c_\alpha$ is a linear character of Δ^* in $\mathfrak{Z}$. By Theorem 2B, there exists a c.c. D_α of Δ^* in $\mathfrak{H}$ such that $\operatorname{tr} D_\alpha\equiv d_\alpha$. Set $A_\alpha=C_\alpha D_\alpha$. Then (Ia) holds and $\operatorname{tr} A_\alpha\equiv a_\alpha$.

Now let B_α stand for the expression on the right of (IIa) with the A_α as already determined. Form the expression

(5.5) $$\frac{B^\beta_\alpha B_\beta}{B_{\alpha\beta}}$$

directly from the definition of B_α. Divide (Ia) by $F_{\alpha,\beta}$, solving for the A-expression, so to speak. Using this in (5.5) one replaces each of the g A-expressions by a quotient of two F's. Each numerator cancels the following denominator, and (5.5) reduces to

$$\frac{F^{\beta}_{\rho\sigma,\alpha}F_{\rho\sigma,\beta}F_{\rho\sigma\alpha\beta\rho^{-\sigma},\rho\sigma}}{F^{\beta}_{\rho\sigma\alpha\rho^{-\sigma},\rho\sigma}F_{\rho\sigma\beta\rho^{-\sigma},\rho\sigma}F_{\rho\sigma,\alpha\beta}}\cdot\frac{F^{\rho\sigma}_{\rho\sigma\alpha\rho^{-\sigma},\rho\sigma\beta\rho^{-\sigma}}}{F_{\alpha,\beta}}. \tag{5.6}$$

The left factor is the original F-part of (5.5); the right factor is all that remains of the A-part. By means of the f.s. identities (5.4), with (α, β, γ) replaced in turn by $(\rho^\sigma, \alpha, \beta)$, $(\rho^\sigma\alpha\rho^{-\sigma}, \rho^\sigma, \beta)$, and $(\rho^\sigma\alpha\rho^{-\sigma}, \rho^\sigma\beta\rho^{-\sigma}, \rho^\sigma)$, (5.6) is found to reduce to 1. Hence B_α is a c.c. of Δ^* in $\mathfrak{H}$.

But (IIb) asserts that B_α has zero trace, since we already have tr $A_\alpha \equiv a_\alpha$. By Theorem 1B, there exists B in $\mathfrak{H}$ such that $B_\alpha = B^{1-\alpha}$. B then satisfies (IIa). It does not necessarily satisfy (IIIa) of course, and our next task is to find B_1 in $\mathfrak{H}$ which satisfies both, i.e.,

$$B_1^{1-\alpha} = B^{1-\alpha} \quad (\text{all } \alpha \in \Delta^*), \qquad B_1^{1-\rho} = A_{\rho\sigma}. \tag{5.7}$$

For $\alpha = \rho^\sigma$, (IIa) gives

$$B^{1-\rho\sigma} = A_{\rho\sigma}^{1+\rho+\cdots+\rho\sigma-1}.$$

Hence

$$N_\rho A_{\rho\sigma} = N_{\rho\sigma}B^{1-\rho^\sigma} = 1. \tag{5.8}$$

By (IIa),

$$B^{(1-\rho)(1-\alpha)} = \frac{B^{1-\alpha}}{B^{(1-\rho\alpha\rho^{-1})\rho}} = \frac{F_{\rho\sigma,\alpha}F^{\rho}_{\rho\sigma+1\alpha\rho-\sigma-1,\rho\sigma}}{F_{\rho\sigma\alpha\rho-\sigma,\rho\sigma}F^{\rho}_{\rho\sigma,\rho\alpha\rho-1}}\cdot\frac{A_\alpha}{A^{\rho\sigma}_{\rho\sigma\alpha\rho-\sigma}}.$$

By (Ia) with (α, β) replaced by (ρ^σ, α) and $(\rho^\sigma\alpha\rho^{-\sigma}, \rho^\sigma)$ this reduces to

$$B^{(1-\rho)(1-\alpha)} = A_{\rho\sigma}^{1-\alpha}.$$

Hence the element $A_{\rho\sigma}B^{-1+\rho}$ is invariant under Δ^*. By (5.8) its ρ-norm is 1. By (IIIb) its trace is zero. Hence by Lemma 2 of §1 there exists an element C invariant under Δ^* such that

$$A_{\rho\sigma}B^{-1+\rho} = C^{1-\rho}.$$

Then $B_1 = BC$ has the desired property (5.7), and so satisfies (IIa) and (IIIa). We have therefore succeeded in finding elements A_α, B of $\mathfrak{H}$ (we now write B for B_1) satisfying (Ia), (IIa), and (IIIa) and such that tr $A_\alpha \equiv a_\alpha$. It remains only to show that they can be modified so as to satisfy the second condition of (5.3), tr $B \equiv b$.

Since A_α and B satisfy (Ia), (IIa), (IIIa), the equations (5.1) define an extension $\mathfrak{G}=\sum u_\sigma \mathfrak{H}$ of $\mathfrak{H}$ by Δ. Let r be the order of ρ; it is divisible by g, say $r=gh$. Raising (5.1c) to the power h,

$$u_\rho^r = u_{\rho\rho}^h N_{\rho\rho} B.$$

By hypothesis for induction, the f.s. $F_{\alpha,\beta}$ of Δ^* in $\mathfrak{H}$ is normalized, and hence $u_{\rho\rho}^h = 1$. Since u_ρ commutes with u_ρ^r, the latter is invariant under ρ and by Lemma 1 must be the ρ-norm of an element D of $\mathfrak{H}$:

$$N_{\rho\rho} B = N_\rho D.$$

Taking traces,

$$h \operatorname{tr} B \equiv hg \operatorname{tr} D.$$

Since n is divisible by hg it follows that $\operatorname{tr} B$ is divisible by g, say

$$\operatorname{tr} B \equiv g k_1.$$

By (IVb), b is also divisible by g, say

$$b \equiv g k_2.$$

Now let E be any element of $\mathfrak{H}$ such that

$$\operatorname{tr} E \equiv k_2 - k_1,$$

and let

$$A'_\alpha = A_\alpha E^{1-\alpha}.$$

Since $E^{1-\alpha}$ is a c.c., (Ia) will remain valid with A'_α instead of A_α. If we replace A_α by A'_α in the right-hand member of (IIa), the new expression is equal to the old one multiplied by E with the exponent

$$\begin{aligned}(1-\alpha) + (1-\rho\alpha\rho^{-1})\rho + \cdots + (1-\rho^{\sigma-1}\alpha\rho^{-\sigma+1})\rho^{\sigma-1} \\ = (1+\rho+\cdots+\rho^{\sigma-1})(1-\alpha).\end{aligned}$$

Hence (IIa) will remain valid if we replace B by

$$B' = E^{1+\rho+\cdots+\rho^{\sigma-1}} B.$$

(IIIa) will also remain valid, since

$$B'^{1-\rho} = E^{1-\rho^\sigma} B^{1-\rho} = E^{1-\rho^\sigma} A_{\rho\sigma} = A'_{\rho\sigma}.$$

But we now have

$$\operatorname{tr} A'_\alpha \equiv \operatorname{tr} A_\alpha \equiv a_\alpha,$$

$$\operatorname{tr} B' \equiv g \operatorname{tr} E + \operatorname{tr} B \equiv g(k_2 - k_1) + g k_1 \equiv b.$$

Hence the elements A'_σ, B' satisfy (Ia), (IIa), (IIIa) and the trace conditions (5.3), which concludes the proof of the theorem.

6. **The subgroup of vanishing traces.** Let $\mathfrak{H}_0$ be the subgroup of $\mathfrak{H}$ consisting of all A in $\mathfrak{H}$ with $\operatorname{tr} A \equiv 0$. $\mathfrak{H}/\mathfrak{H}_0$ is evidently a cyclic group of order n. Let Γ^* be a solvable subgroup of Γ of order m. We conclude the paper by proving the following two statements.

(1) *The multiplicator of* Γ^* *in* $\mathfrak{H}_0$ *is* 1.

(2) *The crossed character group of* Γ^* *in* $\mathfrak{H}_0$ *is cyclic of order* m.

Let $F_{\sigma,\tau}$ be an f.s. of Γ^* in $\mathfrak{H}_0$. Since this means that $\operatorname{tr} F_{\sigma,\tau} \equiv 0$, it follows from Theorem 1A that elements A_σ exist in $\mathfrak{H}$ such that

$$F_{\sigma,\tau} = \frac{A_\sigma^\tau A_\tau}{A_{\sigma\tau}}.$$

Let $a_\sigma = \operatorname{tr} A_\sigma$. Taking traces, we obtain $0 \equiv a_\sigma + a_\tau - a_{\sigma\tau}$. By Theorem 2B, there exists a c.c. B_σ of Γ^* in $\mathfrak{H}$ such that $\operatorname{tr} B_\sigma \equiv a_\sigma$. Setting $C_\sigma = A_\sigma B_\sigma^{-1}$ we have

$$F_{\sigma,\tau} = \frac{C_\sigma^\tau C_\tau}{C_{\sigma\tau}}$$

with $\operatorname{tr} C_\sigma \equiv 0$. Hence $F_{\sigma,\tau}$ is associate to 1 in $\mathfrak{H}_0$, proving (1).

Let A_σ be a c.c. of Γ^* in $\mathfrak{H}$. Since this means that $\operatorname{tr} A_\sigma \equiv 0$, it follows from Theorem 1B that A_σ is a unit c.c. of Γ^* in $\mathfrak{H}$. Conversely every unit c.c. $A^{1-\sigma}$ of Γ^* in $\mathfrak{H}$ lies in $\mathfrak{H}_0$. Hence the group $U\mathfrak{H}_0$ of c.c.'s of Γ^* in $\mathfrak{H}_0$ coincides with the group $\mathfrak{H}^{1-\sigma}$ of unit c.c.'s of Γ^* in $\mathfrak{H}$, and the crossed character group $C\mathfrak{H}_0$ of Γ^* in $\mathfrak{H}_0$ is by definition $\mathfrak{H}^{1-\sigma}/\mathfrak{H}_0^{1-\sigma}$.

Let $N_{\Gamma^*}\mathfrak{H}$ be the subgroup of $\mathfrak{H}$ invariant under Γ^*. Considering the homomorphism $A \to A^{1-\sigma}$ of $\mathfrak{H}$ onto $\mathfrak{H}^{1-\sigma}$ we see that $\mathfrak{H}^{1-\sigma} \simeq \mathfrak{H}/N_{\Gamma^*}\mathfrak{H}$. The subgroup $N_{\Gamma^*}\mathfrak{H}_0$ of $\mathfrak{H}_0$ invariant under Γ^* is evidently the intersection $\mathfrak{H}_0 \cap N_{\Gamma^*}\mathfrak{H}$. (That it coincides with the group of Γ^*-norms of elements of $\mathfrak{H}_0$ follows from the corollary to Lemma 3, but we do not need this.) Hence $\mathfrak{H}_0^{1-\sigma} \simeq \mathfrak{H}_0/N_{\Gamma^*}\mathfrak{H}_0 \simeq \mathfrak{H}_{\Gamma^*}/N_{\Gamma^*}\mathfrak{H}$, where $\mathfrak{H}_{\Gamma^*}$ is the group generated by $\mathfrak{H}_0$ and $N_{\Gamma^*}\mathfrak{H}$. Since $N_{\Gamma^*}\mathfrak{H}$ is the group of Γ^*-norms of elements of $\mathfrak{H}$ (Lemma 1), the trace of every element thereof, and hence the trace of every element of $\mathfrak{H}_{\Gamma^*}$, is divisible by m. Conversely, if the trace of A in $\mathfrak{H}$ is divisible by m, say $\operatorname{tr} A \equiv lm$, and we select any element B of trace l, then $C = N_{\Gamma^*}B$ has trace lm and $A = AC^{-1} \cdot C$ is in $\mathfrak{H}_{\Gamma^*}$; for AC^{-1} is in $\mathfrak{H}_0$ and C is in $N_{\Gamma^*}\mathfrak{H}$. Hence $\mathfrak{H}_{\Gamma^*}$ consists of all elements of $\mathfrak{H}$ whose traces are divisible by m. (2) then follows from $\mathfrak{H}^{1-\sigma}/\mathfrak{H}_0^{1-\sigma} \simeq \mathfrak{H}/\mathfrak{H}_{\Gamma^*}$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASS.
HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON THE BEHAVIOR OF TRIGONOMETRIC SERIES AND POWER SERIES

BY

J. MARCINKIEWICZ AND A. ZYGMUND

INTRODUCTION

1.1. The present paper consists of two parts. The first part contains proofs of theorems describing the behavior of the partial sums of trigonometric series. The second part gives corresponding results for the partial sums of power series on the circle of convergence.

Let us consider the trigonometric series

$$\frac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta) \tag{1.1.1}$$

and the conjugate series

$$\sum_{\nu=1}^{\infty} (a_\nu \sin \nu\theta - b_\nu \cos \nu\theta). \tag{1.1.2}$$

The partial sums of these two series will be denoted by $s_n(\theta)$ and $\tilde{s}_n(\theta)$ respectively. The Cesàro means of order α of the series (1.1.1) and (1.1.2) will be denoted by $\sigma_n^\alpha(\theta)$ and $\tilde{\sigma}_n^\alpha(\theta)$[1]. We also introduce the following notation:

$$\liminf_{n\to\infty} \sigma_n^\alpha(\theta) = \sigma_\alpha(\theta), \qquad \limsup_{n\to\infty} \sigma_n^\alpha(\theta) = \sigma^\alpha(\theta), \qquad \sigma^\alpha(\theta) - \sigma_\alpha(\theta) = \omega_\alpha(\theta).$$

The last expression represents the oscillation of the sequence $\{\sigma_n^\alpha(\theta)\}$ for $n\to+\infty$. If $\sigma_n^\alpha(\theta)$ tends to $+\infty$ or to $-\infty$, we write $\omega_\alpha(\theta)=+\infty$. Hence the condition $\omega_\alpha(\theta)=0$ is both necessary and sufficient for the series (1.1.1) to be summable (C, α) at the point θ.

Similarly, we write

$$\liminf_{n\to\infty} \tilde{\sigma}_n^\alpha(\theta) = \tilde{\sigma}_\alpha(\theta), \qquad \limsup_{n\to\infty} \tilde{\sigma}_n^\alpha(\theta) = \tilde{\sigma}^\alpha(\theta), \qquad \tilde{\sigma}^\alpha(\theta) - \tilde{\sigma}_\alpha(\theta) = \tilde{\omega}_\alpha(\theta).$$

Theorems 1 and 2 which follow connect the behavior of the series (1.1.1) and (1.1.2). By summability A we mean Abel summability.

THEOREM 1. *Suppose that the series* (1.1.1) *is summable A to sum* $s(\theta)$ *at every point of a set E of positive measure, and that for an* $\alpha > -1$

Presented to the Society, January 1, 1941; received by the editors October 10, 1940.

(1) In §1.5 below a few properties of the arithmetic means are collected which will be used in the sequel.

$$\sigma_\alpha(\theta) > -\infty \qquad \text{for } \theta \in E. \tag{1.1.3}$$

Then at almost every point of E we have the following relations:

$$\sigma^\alpha(\theta) < +\infty, \tag{1.1.4}$$

$$-\infty < \bar{\sigma}_\alpha(\theta) \leqq \bar{\sigma}^\alpha(\theta) < +\infty, \tag{1.1.5}$$

$$\tilde{\omega}_\alpha(\theta) = \omega_\alpha(\theta). \tag{1.1.6}$$

Moreover the series (1.1.1) *and* (1.1.2) *are summable* $(C, \alpha+\epsilon)$, $\epsilon>0$, *at almost every point of the set E, and*

$$s(\theta) = \tfrac{1}{2}\{\sigma_\alpha(\theta) + \sigma^\alpha(\theta)\}, \tag{1.1.7}$$

$$\tilde{s}(\theta) = \tfrac{1}{2}\{\tilde{\sigma}_\alpha(\theta) + \tilde{\sigma}^\alpha(\theta)\} \tag{1.1.8}$$

almost everywhere in E, where $\tilde{s}(\theta)$ denotes the $(C, \alpha+\epsilon)$ sum of the series (1.1.2).

If the hypotheses of Theorem 1 are satisfied, we have therefore that the αth Cesàro means of the series (1.1.1) and (1.1.2) are bounded at almost every point of E.

From (1.1.6) we see that the oscillations of the sequences $\{\sigma_n^\alpha(\theta)\}$ and $\{\tilde{\sigma}_n^\alpha(\theta)\}$ are equal at almost every point of the set E. In particular, if one of the series (1.1.1) and (1.1.2) is summable (C, α) at almost every point of E, so is the other.

The relations (1.1.7) and (1.1.8) show that the limits of indetermination of the sequences $\{\sigma_n^\alpha(\theta)\}$ and $\{\tilde{\sigma}_n^\alpha(\theta)\}$ are situated symmetrically with respect to the (generalized) sums of the series (1.1.1) and (1.1.2).

It follows at once from Theorem 1 that, if the series (1.1.1) is summable A at every point of the set E, and if $\sigma_\alpha(\theta) = -\infty$ in E, then $\sigma^\alpha(\theta) = +\infty$ almost everywhere in E. The question whether we also have

$$\bar{\sigma}_\alpha(\theta) = -\infty, \quad \bar{\sigma}^\alpha(\theta) = +\infty$$

almost everywhere in E is left open.

THEOREM 2. *Suppose that the αth Cesàro means (where $\alpha > -1$) of the series* (1.1.1) *are bounded at every point of a set E of positive measure, that is, that*

$$\sigma_n^\alpha(\theta) = O(1) \qquad \text{for every } \theta \in E. \tag{1.1.9}$$

Then at almost every point of E the series (1.1.1) *and* (1.1.2) *are summable* $(C, \alpha+\epsilon)$ $(\epsilon>0)$ *and we have the relations* (1.1.5), (1.1.6), (1.1.7) *and* (1.1.8).

The hypothesis (1.1.3) of one-sided boundedness, in Theorem 1, is replaced in Theorem 2 by the stronger hypothesis (1.1.9); but, on the other hand, no assumption is made in Theorem 2 concerning the summability of the series (1.1.1).

Theorems 1 and 2 are known. Special cases of Theorem 1 have been proved

by Kuttner(2) and Plessner(3), and the general result stated above was proved by the authors(4).

The original proofs of Theorems 1 and 2 are however rather difficult. In the first part of this paper we give new proofs of these theorems, using certain ideas of Plessner, in particular the application of summability (C^*, α) (defined in §2.1 below(5)). The chief simplification in comparison with the earlier proofs (in particular with Plessner's proof) is that we do not require the theory of higher generalized derivatives.

1.2. Theorems 1 and 2 raise the problem of the behavior of the Cesàro means of a power series on the circle of convergence. Let us begin first with the case of partial sums.

Let

$$F(z) = \sum_{\nu=0}^{\infty} c_\nu z^\nu, \qquad |z| < 1, \tag{1.2.1}$$

be a function regular inside the unit circle. We may suppose for the moment without loss of generality (subtracting a constant from $F(z)$ if necessary) that c_0 is real. If

$$c_n = a_n - ib_n \qquad \text{for } n = 1, 2, 3, \cdots,$$

then the real and imaginary parts of the series

$$\sum_{\nu=0}^{\infty} c_\nu e^{i\nu\theta} \tag{1.2.2}$$

may be written in the forms (1.1.1) and (1.1.2) respectively.

Let us assume that the partial sums of the series (1.2.1) are bounded at every point $z = e^{i\theta}$ of a certain set E situated on the circumference $|z| = 1$. We shall also denote by E the set of the corresponding arguments θ; this, however, will not lead to any confusion. By Theorem 2 the series (1.2.2) is summable $(C, 1)$ to sum $t(\theta)$ at almost every point of E. It follows from Theorem 1, therefore, that for almost every point θ of E the limit points of the sequence

$$t_n(\theta) = \sum_{\nu=0}^{n} c_\nu e^{i\nu\theta}, \qquad n = 0, 1, 2, \cdots,$$

(2) Kuttner [1] proves that if the series (1.1.1) converges in E and the series (1.1.2) is summable $(C, 1)$ in E, then the latter series converges almost everywhere in E.

(3) The result of Plessner [1] coincides with Lemma M of the present paper in the case $\beta = 0, 1, 2, \cdots$. Plessner only sketches the proof of the result.

(4) See Marcinkiewicz and Zygmund [1].

(5) That notion was considered independently by one of the authors and applied to the theory of Fourier series. See Zygmund [1, p. 61, Example 4].

It must however be observed that the expressions

$$\theta_n^\alpha(\theta + h_n) \pm \sigma_n^\alpha(\theta - h_n)$$

which play an essential part in our argument have been studied for the first time by Rogosinski [1].

belong to the square with center $t(\theta)$ and sides $\omega_0(\theta)$, parallel to the axes. Here $\omega_0(\theta)$ denotes the oscillation of the real part of the series (1.2.2). By Theorem 2, this square cannot be replaced by a smaller rectangle with the same center and sides parallel to the axes. On the other hand, not every point of the square need be a limit point of the sequence $\{t_n(\theta)\}$, as simple examples show.

In fact, take the series

$$1 + z + z^2 + \cdots, \tag{1.2.3}$$

the partial sums of which are bounded and summable $(C, 1)$ at every point z of $|z| = 1$ except $z = 1$. Here

$$t(\theta) = \frac{1}{1 - e^{i\theta}}, \qquad t_n(\theta) - t(\theta) = -\frac{e^{i(n+1)\theta}}{1 - e^{i\theta}},$$

so that all the partial sums of the series (1.2.3), with $z = e^{i\theta}$, $\theta \not\equiv 0 \pmod{2\pi}$, are situated on the circle with center $(1 - e^{i\theta})^{-1}$ and radius $|1 - e^{i\theta}|^{-1}$. The same is of course true of the set of the limit points of the sequence $\{t_n(\theta)\}$. Moreover, if θ is not commensurable with π, every point of the circle just defined is a limit of the sequence $\{t_n(\theta)\}$.

Let us modify this example slightly by considering the power series

$$\begin{aligned}(\alpha + \beta) + \alpha z + (\alpha + \beta)z^2 + \alpha z^3 + \cdots \\ = \alpha(1 + z + z^2 + \cdots) + \beta(1 + z^2 + z^4 + \cdots),\end{aligned} \tag{1.2.4}$$

where α and β are any constants. It may be verified without difficulty that in this case, for θ different from 0 and π, the partial sums $t_n(\theta)$ are situated on two concentric circles, with center

$$\frac{\alpha}{1 - e^{i\theta}} + \frac{\beta}{1 - e^{2i\theta}}$$

and radii

$$\frac{1}{|1 - e^{i\theta}|} \cdot \left| \alpha + \frac{\beta}{1 + e^{\pm i\theta}} \right|.$$

These two circles are different if the ratio β/α is not real. If θ is incommensurable with π, every point of these two circles is a limit point of the sequence $\{t_n(\theta)\}$.

Similarly, in the case of the series

$$\alpha(1 + z + z^2 + \cdots) + \beta(1 + z^2 + z^4 + \cdots) + \gamma(1 + z^3 + z^6 + \cdots \tag{1.2.5}$$

the partial sums $t_n(\theta)$ are situated on six (generally distinct) circles with center

$$t(\theta) = \frac{\alpha}{1 - e^{i\theta}} + \frac{\beta}{1 - e^{2i\theta}} + \frac{\gamma}{1 - e^{3i\theta}},$$

and if θ is incommensurable with π, every point of each of those six circles is a limit point of the sequence $\{t_n(\theta)\}$.

A slightly different example may be obtained by considering the series

$$(1 + z + z^2 + \cdots) + \alpha(1 + e^{i\lambda}z + e^{2i\lambda}z^2 + \cdots), \qquad z = e^{i\theta},$$

where $\alpha \neq 0$ and λ is a real number incommensurable with π. If θ is linearly independent of λ, and n is suitably chosen, the fractional parts of $(n+1)\theta$ and $(n+1)\lambda$ may be as close as we wish to any preassigned numbers of the interval $(0, 1)$. Using this fact, we easily obtain that, except for a denumerable set of values of θ, the limit points of the sequence of the partial sums of our series, with $z=e^{i\theta}$, form an annulus not reducing to a circle or a circumference.

By combining this example with the preceding ones, we may obtain a power series such that for values of θ belonging to a set of positive measure the set of limit points of the partial sums will consist of several concentric annuli.

The particular series which we have just considered illustrate some general theorems which will be stated in a moment. It will be convenient to introduce certain notations.

By

$$\mathrm{A}\{z_0; \alpha, \beta\}$$

where $0 \leqq \alpha \leqq \beta \leqq \infty$, we shall mean the annulus consisting of the points z satisfying the inequalities $\alpha \leqq |z - z_0| \leqq \beta$. The circle $\mathrm{A}\{z_0; 0, \beta\}$ will be denoted simply by

$$\mathrm{K}\{z_0; \beta\},$$

and the circumference $|z - z_0| = \beta$ by

$$\mathrm{C}\{z_0; \beta\}.$$

We shall say that a plane set Z is *of circular structure*, if there is a point z_0 (*center* of Z) such that whenever a point ζ belongs to Z so does the whole circumference $\mathrm{C}\{z_0; |\zeta - z_0|\}$.

Given a function $F(z)$ defined by (1.2.1) we shall write

$$M(\theta) = M_F(\theta) = \limsup_{n\to\infty} |t_n(\theta) - t(\theta)|,$$

$$m(\theta) = m_F(\theta) = \liminf_{n\to\infty} |t_n(\theta) - t(\theta)|,$$

where $t(\theta)$ is the $(C, 1)$ sum (provided it exists) of the series (1.2.2). The set of the limit points of the sequence $\{t_n(\theta)\}$, when θ is fixed, will be denoted by $\mathrm{L}_F(\theta)$, or simply by $\mathrm{L}(\theta)$.

Now we may state the following

THEOREM A. *Suppose that the partial sums of the series* (1.2.1) *are bounded at every point of a set E situated on the circumference* $\mathrm{C}\{0; 1\}$ *and let* $t(\theta)$ *denote*

the $(C, 1)$ *sum (existing, by Theorem 2, almost everywhere in* E*) of the series* (1.2.2). *Then, for almost every* $\theta \in E$, *the set* $L(\theta)$ *is of circular structure, with center* $t(\theta)$.

In other words, for almost every value of θ contained in E the set of the limit points of the sequence $\{t_n(\theta)\}$ consists of a finite, denumerable, or non-denumerable set of circumferences with center at $t(\theta)$. The set $L(\theta)$ is obviously contained in the annulus $A\{t(\theta); m(\theta), M(\theta)\}$. The set $L(\theta)$ being closed, the extreme circumferences $C\{t(\theta); m(\theta)\}$ and $C\{t(\theta); M(\theta)\}$ of that annulus belong to $L(\theta)$.

If the partial sums of the series (1.2.2) are bounded for some $\theta = \theta_0$, the coefficients c_n are bounded. The examples of the series (1.2.4) and (1.2.5) show that the set $L(\theta)$ need not coincide with $A\{t(\theta); m(\theta), M(\theta)\}$. This however will be the case, if we assume additionally that the coefficients c_n tend to 0.

In other words, we have the following theorem:

THEOREM B. *Suppose that the series* (1.2.2) *satisfies the conditions of Theorem* A *and that* $c_n \to 0$. *Then*

$$L(\theta) = A\{t(\theta); m(\theta), M(\theta)\}$$

for almost every θ *of* E.

The deduction of this theorem from Theorem A is obvious. Since $t_n(\theta) - t_{n-1}(\theta) \to 0$, every annulus $A\{t(\theta); r_1, r_2\}$, with $m(\theta) < r_1 < r_2 < M(\theta)$, contains infinitely many points $t_n(\theta)$, and so also points of the set $L(\theta)$. For every value of θ for which the set $L(\theta)$ is of circular structure that set must contain all the circumferences $C\{t(\theta); r\}$ with $m(\theta) \leqq r \leqq M(\theta)$. This completes the proof.

Assume that for almost every $\theta \in E$, $m(\theta) = 0$; that is, that for almost every value θ of E there is a sequence $\{n_k\}$ (which may depend on θ) such that

$$t_{n_k}(\theta) \to t(\theta). \tag{1.2.6}$$

Let us suppose furthermore that c_n tends to 0. Then obviously

$$L(\theta) = K\{t(\theta); M(\theta)\}$$

almost everywhere in E. This is certainly the case if the real part of the series (1.2.2) is a Fourier-Lebesgue series. For then, as is well known,

$$\int_0^{2\pi} |t(\theta) - t_n(\theta)|^{\mu} d\theta \to 0 \qquad \text{for every } 0 < \mu < 1(^6),$$

and so there is a sequence $\{n_k\}$, independent of θ, such that (1.2.6) is satisfied almost everywhere.

(6) See, e.g., Zygmund [1, p. 153].

1.3. The results stated in the preceding section for the partial sums of power series can be extended to the arithmetic means of power series(7).

Let $\tau_n^\alpha(\theta)$ denote the αth Cesàro means of the series (1.2.2). By

$$L_P^\alpha(\theta) = L^\alpha(\theta)$$

we shall denote the set of the limit points of the sequence $\{\tau_n^\alpha(\theta)\}$.

THEOREM 3. *Suppose that the series* (1.2.2) *is summable* $(C, \alpha+1)$ *(where* $\alpha > -1$*) at every point* θ *of a set* E*, to sum* $t(\theta)$*. Then at almost every point* θ *of* E *the set* $L^\alpha(\theta)$ *is of circular structure, with center* $t(\theta)$*.*

In particular, the set $L^\alpha(\theta)$ *is of circular structure almost everywhere in* E*, if the arithmetic means* $\tau_n^\alpha(\theta)$ *of the series* (1.2.2) *are bounded at every point* θ *of* E*.*

That the second part of this theorem is a consequence of the first part follows at once from Theorem 2.

1.4. THEOREM 4. *Suppose that the conditions of Theorem 3 are satisfied for some* $\alpha \geqq 0$ *and that* $c_n \to 0$*. Let*

$$(1.4.1) \qquad m^\alpha(\theta) = \liminf_{n\to\infty} |\tau_n^\alpha(\theta) - t(\theta)|, \quad M^\alpha(\theta) = \limsup_{n\to\infty} |\tau_n^\alpha(\theta) - t(\theta)|.$$

Then

$$L^\alpha(\theta) = A\{t(\theta); m(\theta), M(\theta)\}$$

almost everywhere in E*.*

This theorem follows in the same way from the preceding theorem as Theorem B follows from Theorem A, provided we prove that

$$(1.4.2) \qquad \tau_n^\alpha(\theta) - \tau_{n-1}^\alpha(\theta) \to 0$$

which is immediate (cf. §5.3).

In the remainder of Part II we give a proof of a theorem (Theorem 5 below) which was already stated without proof in an earlier paper, and indicate extensions of the previous results to the case of Dirichlet series.

1.5. In this section we state without proofs a few known results from the theory of Cesàro arithmetic means. We state only the facts we shall actually need in the sequel.

Given a series

$$(1.5.1) \qquad \sum_{n=0}^{\infty} u_n$$

(7) For the theory of Cesàro means, see Andersen [1] and Kogbetliantz [1]. All the results required for the present paper will be found in Zygmund [1].

and a number α, we define numbers s_n^α $(n=0, 1, \cdots)$ by the equation

$$\sum_{n=0}^{\infty} s_n^\alpha x^n = \frac{1}{(1-x)^{\alpha+1}} \sum_{n=0}^{\infty} u_n x^n.$$

Let us consider the numbers A_n^α defined by the formula

$$\sum_{n=0}^{\infty} A_n^\alpha x^n = \frac{1}{(1-x)^{\alpha+1}}.$$

If the ratio

$$\sigma_n^\alpha = s_n^\alpha / A_n^\alpha \tag{1.5.2}$$

tends to s as n tends to infinity, the series (1.5.1) is said to be summable by the Cesàro method of arithmetic means of order α, or simply summable (C, α), to sum s. The numbers s_n^α and σ_n^α are called respectively the Cesàro sums and the Cesàro means of order α of the series (1.5.1). If the ratio (1.5.2) is bounded [bounded below] as $n \to \infty$, the series (1.5.1) is said to be bounded (C, α) [bounded below (C, α)].

From the above definitions one can easily deduce the following properties of Cesàro means.

(i) If the series (1.5.1) is summable (C, α), where $\alpha > -1$, to sum s, then the series is also summable (C, β) to sum s, provided that $\beta > \alpha$.

If the series (1.5.1) is bounded [bounded below] (C, α), then it is also bounded [bounded below] (C, β) for $\beta > \alpha > -1$.

(ii) If the series (1.5.1) is summable (C, α), $\alpha > -1$, to sum s, then it is also summable by Abel's method (summable A) to s. In other words

$$\lim_{z \to 1-0} \sum_{n=0}^{\infty} u_n z^n = s.$$

Here the variable z tends to 1 along the real axis, but the result holds if z tends to 1 from inside the circle $|z| < 1$ along any non-tangential path, that is, along any path contained between two chords of the circumference $|z| = 1$ passing through the point $z = 1$.

(iii) If the series (1.5.1) is summable (C, α), $\alpha > -1$, then $u_n = o(n^\alpha)$.

(iv) We have

$$A_n^\alpha = \frac{(\alpha+1)(\alpha+2)\cdots(\alpha+n)}{n!} \simeq \frac{n^\alpha}{\Gamma(\alpha+1)},$$

where the relation $a_n \simeq b_n$ means that $a_n / b_n \to 1$.

In particular, the numbers A_n^α are non-negative for $\alpha > -1$, form an increasing sequence tending to $+\infty$ when $\alpha > 0$, and decrease to 0 when $-1 < \alpha < 0$. Moreover $A_n^0 = 1$ for all n.

(v) The numbers A_n^α and s_n^α satisfy the following relations

$$\sum_{\nu=0}^{n} A_\nu^\alpha A_{n-\nu}^\beta = A_n^{\alpha+\beta+1}, \qquad \sum_{\nu=0}^{n} s_\nu^\alpha A_{n-\nu}^\beta = s_n^{\alpha+\beta+1}. \tag{1.5.3}$$

In particular,

$$\sum_{\nu=0}^{n} A_\nu^\alpha = A_n^{\alpha+1}, \qquad \sum_{\nu=0}^{n} s_\nu^\alpha = s_n^{\alpha+1},$$

and so

$$A_n^{\alpha+1} - A_{n-1}^{\alpha+1} = A_n^\alpha, \qquad s_n^{\alpha+1} - s_{n-1}^{\alpha+1} = s_n^\alpha.$$

(vi) If $\{\epsilon_\nu\}$ is any sequence tending to 0, then

$$\sum_{\nu=0}^{n} \epsilon_\nu A_\nu^\alpha A_{n-\nu}^\beta = o(A_n^{\alpha+\beta+1}) \qquad \text{for } \alpha > -1,\ \beta > -1. \tag{1.5.4}$$

This follows from (i) and from the second formula (1.5.3), if we put $\epsilon_\nu A_\nu^\alpha = s_\nu^\alpha$, and observe that the series summable (C, α) to 0 is also summable $(C, \alpha+\beta+1)$ to 0.

(vii) It is easy to see that (1.5.4) holds if we replace the upper limit of summation n by $n-k$, where k is any fixed positive integer. Hence

$$\sum_{\nu=0}^{n-k} o(A_\nu^\alpha) A_{n-\nu}^\beta = o(A_n^{\alpha+\beta+1}) \qquad \text{for } \alpha, \beta > -1,\ k = \text{const.} \tag{1.5.5}$$

PART I

2. LEMMAS ON SUMMABILITY (C^*, α)

2.1. Let $\sigma_n^\alpha(\theta)$ denote the αth Cesàro means of the series

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta). \tag{2.1.1}$$

We shall say that this series is summable (C^*, α) at the point θ_0 to sum s, if

$$\sigma_n^\alpha(\theta_0 + h_n) \to s$$

for every sequence $h_n = O(1/n)$. This property is equivalent to the following one: to every constant $A > 0$ and to every $\epsilon > 0$ corresponds an integer $n_0 = n_0(A, \epsilon)$ such that

$$|\sigma_n^\alpha(\theta_0 + h) - s| \leqq \epsilon \qquad \text{for } |h| \leqq A/n,\ n > n_0. \tag{2.1.2}$$

LEMMA A. *A necessary and sufficient condition that the series*

$$\sum_{\nu=0}^{\infty} a_\nu \tag{2.1.3}$$

should be summable (C, α), *where* $\alpha > -1$, *to sum* s *is that the series*

$$\sum_{\nu=0}^{\infty} a_\nu \cos \nu\theta \tag{2.1.4}$$

should be summable (C^*, α) *at the point* 0 *to sum* s[8].

Proof. The sufficiency of the condition is evident. For the proof of the necessity we may assume that $s=0$. We begin with the case of integral α, so that $\alpha=0, 1, 2, \cdots$.

We write

$$\mu_\nu = e^{i\nu\theta}, \qquad \epsilon_\nu = \epsilon_\nu^{(n)} = A_{n-\nu}^{\alpha}\mu_\nu.$$

Let t_n^α denote the αth Cesàro sums of the series (2.1.3), and let $\sigma_n^\alpha(\theta)$ be the αth Cesàro means of the series

$$\sum_{\nu=0}^{\infty} a_\nu e^{i\nu\theta}. \tag{2.1.5}$$

It is sufficient to prove that, if $t_n^\alpha = o(n^\alpha)$, then

$$\sigma_n^\alpha(\theta) = o(1) \qquad \text{for } |n\theta| \leqq A. \tag{2.1.6}$$

Applying Abel's transformation $\alpha+1$ times, we may write

$$A_n^\alpha \sigma_n^\alpha(\theta) = \sum_{\nu=0}^{n} a_\nu \epsilon_\nu = \sum_{\nu=0}^{n-\alpha-1} t_\nu^\alpha \Delta^{\alpha+1}\epsilon_\nu + \sum_{k=0}^{\alpha} t_{n-k}^k \Delta^k \epsilon_{n-k} = P_n + Q_n[9], \tag{2.1.7}$$

say. The numbers μ_ν being bounded, the expressions $\Delta^k\epsilon_{n-k}$ are also bounded for $k=0, 1, \cdots, \alpha$. Since the relation $t_n^\alpha = o(n^\alpha)$ implies

$$t_{n-k}^k = o(n^\alpha) \quad \text{for } n \to \infty, \text{ and } k = 0, 1, \cdots, \alpha,$$

it follows that

$$Q_n = o(n^\alpha).$$

In order to estimate the expression P_n, we use the formula

$$\Delta^l \xi_n \eta_n = \sum_{j=0}^{l} C_{l,j} \Delta^j \xi_n \Delta^{l-j} \eta_{n+j}, \qquad l = 0, 1, \cdots. \tag{2.1.8}$$

Hence, taking into account that

---

[8] Lemma A of the present paper is a special case of Lemma D. The latter lemma is stated in Plessner [1], for $\alpha=0, 1, 2, \cdots$.

[9] We use the notation

$$\Delta c_\nu = c_\nu - c_{\nu+1}, \qquad \Delta^k c_\nu = \Delta(\Delta^{k-1} c_\nu).$$

$$\Delta\mu_\nu = e^{i\nu\theta} - e^{i(\nu+1)\theta} = e^{i(\nu+1/2)\theta}\cdot(-2i\sin\tfrac{1}{2}\theta),$$

and generally

$$\Delta^j\mu_\nu = e^{i(\nu+j/2)\theta}(-2i\sin\tfrac{1}{2}\theta)^j,$$

we obtain

$$\begin{aligned}\Delta^{\alpha+1}\epsilon_\nu &= \sum_{j=0}^{\alpha+1} C_{\alpha+1,j}\Delta_\nu^j(A_{n-\nu}^\alpha)\Delta^{\alpha+1-j}\mu_{\nu+j}\\ &= \sum_{j=0}^{\alpha} C_{\alpha+1,j}A_{n-\nu}^{\alpha-j}(-2i\sin\tfrac{1}{2}\theta)^{\alpha+1-j}e^{i(\alpha+1-j)\theta/2}e^{i\nu\theta}.\end{aligned}$$

It follows that P_n is equal to the sum of the expressions

$$e^{i(\alpha+1-j)\theta/2}(-2i\sin\tfrac{1}{2}\theta)^{\alpha+1-j}C_{\alpha+1,j}\sum_{\nu=0}^{n-\alpha-1} t_\nu^\alpha A_{n-\nu}^{\alpha-j}e^{i\nu\theta} \tag{2.1.9}$$

for $j=0, 1, \cdots, \alpha$. The absolute value of (2.1.9) does not exceed (cf. (1.5.5))

$$|\theta|^{\alpha+1-j}C_{\alpha+1,j}\sum_{\nu=0}^{n-\alpha-1} o(A_\nu^\alpha)A_{n-\nu}^{\alpha-j} = |\theta|^{\alpha+1-j}o(n^{2\alpha-j+1}) = o(n^\alpha),$$

provided that $|n\theta| \leqq A$. Hence $P_n=o(n^\alpha)$, and since the same relation was obtained for Q_n, $P_n+Q_n=o(n^\alpha)$. This and (2.1.7) give (2.1.6), so that the lemma is established in the case of integral α.

2.2. In the case of fractional $\alpha>-1$, we write

$$\beta = [\alpha]+1, \quad \mu_\nu = e^{i\nu\theta}, \quad \eta_\nu = \eta_\nu^{(n)}(\theta) = A_{n-\nu}^\alpha(\mu_\nu - \mu_n).$$

Hence β is a non-negative integer greater than α. We write

$$A_n^\alpha\sigma_n^\alpha(\theta) = \mu_n\sum_{\nu=0}^{n} A_{n-\nu}^\alpha a_\nu + \sum_{\nu=0}^{n} a_\nu\eta_\nu = o(n^\alpha) + \sum_{\nu=0}^{n} a_\nu\eta_\nu. \tag{2.2.1}$$

To the last sum we apply Abel's transformation $\beta+1$ times:

$$\sum_{\nu=0}^{n} a_\nu\eta_\nu = \sum_{\nu=0}^{n-\beta-1} t_\nu^\beta\Delta^{\beta+1}\eta_\nu + \sum_{k=1}^{\beta} t_{n-k}^k\Delta^k\eta_{n-k} = P_n + Q_n, \tag{2.2.2}$$

say. Now

$$\begin{aligned}\Delta^k\eta_\nu &= \sum_{j=0}^{k-1} C_{k,j}A_{n-\nu}^{\alpha-j}\Delta^{k-j}\mu_{\nu+j} + A_{n-\nu}^{\alpha-k}(\mu_{\nu+k}-\mu_n)\\ &= \sum_{j=0}^{k-1} C_{k,j}A_{n-\nu}^{\alpha-j}\Delta^{k-j}\mu_{\nu+j} + O(|\theta|)\end{aligned} \tag{2.2.3}$$

for $\nu=n-k$ and $k=1, 2, \cdots, \beta$. Since the series (2.1.3) is summable (C, α) and so also (C, β) to 0, we obtain

$$t_{n-k}^{k} = o(n^{\beta}) \qquad \text{for } k = 0, 1, \cdots.$$

Hence

$$Q_n = \sum_{k=1}^{\beta} o(n^{\beta}) O(|\theta|) = o(n^{\beta-1}) = o(n^{\alpha}) \tag{2.2.4}$$

for $|n\theta| \leqq A$.

It remains to estimate the sum P_n. On account of the *first* equation (2.2.3) with $k=\beta+1$, P_n is a sum of $\beta+1$ expressions

$$C_{\beta+1,j} \sum_{\nu=0}^{n-\beta-1} t_{\nu}^{\beta} A_{n-\nu}^{\alpha-j} \Delta^{\beta+1-j} \mu_{\nu+j}, \qquad j = 0, 1, \cdots, \beta, \tag{2.2.5}$$

and of the expression

$$\sum_{\nu=0}^{n-\beta-1} t_{\nu}^{\beta} A_{n-\nu}^{\alpha-\beta-1} (\mu_{\nu+\beta+1} - \mu_n). \tag{2.2.6}$$

The sum (2.2.5) is

$$O(|\theta|^{\beta+1-j}) \sum_{\nu=0}^{n-\beta-1} o(A_{\nu}^{\beta}) A_{n-\nu}^{\alpha-j} = O(n^{j-\beta-1}) o(n^{\alpha+\beta-j+1}) = o(n^{\alpha}),$$

and the sum (2.2.6) does not exceed in absolute value

$$\sum_{\nu=0}^{n-\beta-1} o(A_{\nu}^{\beta}) A_{n-\nu}^{\alpha-\beta-1} (n-\nu) |\theta| = O(|\theta|) \sum_{\nu=0}^{n-\beta-1} o(A_{\nu}^{\beta}) A_{n-\nu}^{\alpha-\beta}$$
$$= O(|\theta|) o(n^{\alpha+1}) = o(n^{\alpha}).$$

It follows that

$$P_n = o(n^{\alpha}). \tag{2.2.7}$$

From (2.2.1), (2.2.2), (2.2.4), (2.2.7) we deduce (2.1.6), and so the proof of Lemma A is complete.

2.3. LEMMA B. *A necessary and sufficient condition that the series*

$$\sum_{\nu=0}^{\infty} b_{\nu} \sin \nu\theta \tag{2.3.1}$$

should be summable (C^*, α), $\alpha > -1$, *at the point* $\theta=0$, *is that the sequence* $\{\nu b_{\nu}\}$ $(\nu=0, 1, 2, \cdots)$ *should be summable* $(C, \alpha+1)$ *to* 0.

Proof. The argument is similar to that of Lemma A. We begin by proving the sufficiency of the condition, and we assume first that α is a non-negative integer. Let

$$(2.3.2)\qquad \nu b_\nu = c_\nu, \qquad \mu_\nu = \mu_\nu^{(n)}(\theta) = \frac{\sin \nu\theta}{\nu\theta}, \qquad \gamma_\nu = \gamma_\nu^{(n)}(\theta) = A_{n-\nu}^{\alpha}\mu_\nu.$$

By $\sigma_n^\alpha(\theta)$ we shall denote the αth Cesàro means of the series (2.3.1) and by u_n^k the iterated sums of the sequence $\{c_\nu\}$, so that

$$u_n = u_n^0 = c_n, \qquad u_n^k = u_0^{k-1} + u_1^{k-1} + \cdots + u_n^{k-1},$$
$$n = 0, 1, \cdots; \; k = 1, 2, \cdots.$$

Hence

$$(2.3.3)\qquad A_n^\alpha \sigma_n^\alpha(\theta) = \theta \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha} \nu b_\nu \frac{\sin \nu\theta}{\nu\theta} = \theta \sum_{\nu=0}^{n} c_\nu \gamma_\nu.$$

Applying Abel's transformation $\alpha+1$ times we have

$$(2.3.4)\qquad \theta \sum_{\nu=0}^{n} c_\nu\gamma_\nu = \theta \sum_{\nu=0}^{n-\alpha-1} u_\nu^{\alpha+1}\Delta^{\alpha+1}\gamma_\nu + \theta \sum_{k=0}^{\alpha} u_{n-k}^{k+1}\Delta^k\gamma_{n-k} = \theta P_n + \theta Q_n,$$

say. By hypothesis, $u_n^{\alpha+1} = o(n^{\alpha+1})$, and so also $u_n^{k+1} = o(n^{\alpha+1})$ for $k = 0, 1, \cdots, \alpha$. This gives

$$(2.3.5)\qquad \theta Q_n = \theta o(n^{\alpha+1}) = o(n^\alpha),$$

if only $|n\theta| \leqq A$. Since

$$\Delta^{\alpha+1}\gamma_\nu = \sum_{j=0}^{\alpha+1} C_{\alpha+1,j}\Delta_\nu^j A_{n-\nu}^{\alpha}\Delta^{\alpha+1-j}\mu_{\nu+j} = \sum_{j=0}^{\alpha} C_{\alpha+1,j} A_{n-\nu}^{\alpha-j} O(|\theta|^{\alpha+1-j})(10),$$

θP_n is a sum of $\alpha+1$ terms of the form

$$O(|\theta|^{\alpha+2-j}) \sum_{\nu=0}^{n-\alpha-1} o(A_\nu^{\alpha+1}) A_{n-\nu}^{\alpha-j} = O(|\theta|^{\alpha+2-j}) o(n^{2\alpha-j+2}),$$

where $j = 0, 1, \cdots, \alpha$. Hence, if $|n\theta| \leqq A$,

$$(2.3.6)\qquad \theta P_n = o(n^\alpha).$$

From (2.3.3), (2.3.4), (2.3.5) and (2.3.6) we see that

$$(2.3.7)\qquad \sigma_n^\alpha(\theta) \to 0 \qquad \text{for } n \to \infty \text{ and } |n\theta| \leqq A.$$

This completes the proof of sufficiency in Lemma B for the case of integral α.

2.4. If α is fractional, we write

---

(10) Here (and in the sequel) we use the known fact that if a function $\phi(u)$ has a pth derivative, and if $\phi(\nu) = \phi_\nu$, then

$$\Delta^p\phi_\nu = (-1)^p\phi(\nu + \theta p)$$

where θ is a number contained between 0 and 1.

$$\beta = [\alpha] + 1, \qquad \xi_\nu = \xi_\nu^{(n)}(\theta) = A_{n-\nu}^{\alpha}(\mu_\nu - \mu_n),$$

where μ_ν is defined by (2.3.2). Then

$$A_n^\alpha \sigma_n^\alpha(\theta) = \theta\mu_n \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha} c_\nu + \theta \sum_{\nu=0}^{n} c_\nu \xi_\nu.$$

First

$$\theta\mu_n \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha} c_\nu = \theta\mu_n o(n^{\alpha+1}) = o(n^\alpha) \tag{2.4.1}$$

for $|n\theta| \leqq A$. Now

$$\theta \sum_{\nu=0}^{n} c_\nu \xi_\nu = \theta \sum_{\nu=0}^{n-\beta-1} u_\nu^{\beta+1} \Delta^{\beta+1} \xi_\nu + \theta \sum_{k=1}^{\beta} u_{n-k}^{k+1} \Delta^k \mu_{n-k} = \theta P_n + \theta Q_n, \tag{2.4.2}$$

say. But

$$\begin{aligned} \Delta^k \xi_\nu &= \sum_{j=0}^{k-1} C_{k,j} A_{n-\nu}^{\alpha-j} \Delta^{k-j} \mu_{\nu+j} + A_{n-\nu}^{\alpha-k}(\mu_{\nu+k} - \mu_n) \\ &= \sum_{j=0}^{k-1} C_{k,j} A_{n-\nu}^{\alpha-j} \Delta^{k-j} \mu_{\nu+j} = O(|\theta|) \end{aligned} \tag{2.4.3}$$

for $k = 1, 2, \cdots, \beta$ and $\nu = n-k$. By hypothesis, the sequence $\{c_\nu\}$ is summable $(C, \alpha+1)$, and so also $(C, \beta+1)$, to 0. Hence $u_{n-k}^{k+1} = o(n^{\beta+1})$ for $k = 1, 2, \cdots, \beta$. This gives

$$\theta Q_n = \theta \cdot O(|\theta|) o(n^{\beta+1}) = o(n^{\beta-1}) = o(n^\alpha) \tag{2.4.4}$$

for $|n\theta| \leqq A$.

Substituting into the formula for P_n the first formula (2.4.3) with $k = \beta+1$, we see that θP_n is a sum of $\beta+1$ expressions

$$\theta C_{\beta+1,j} \sum_{\nu=0}^{n-\beta-1} o(A_\nu^{\beta+1}) A_{n-\nu}^{\alpha-j} \Delta^{\beta+1-j} \mu_{\nu+j}, \qquad j = 0, 1, \cdots, \beta, \tag{2.4.5}$$

and of the expression

$$\theta \sum_{\nu=0}^{n-\beta-1} o(A_\nu^{\beta+1}) A_{n-\nu}^{\alpha-\beta-1}(\mu_{\nu+\beta+1} - \mu_n). \tag{2.4.6}$$

The absolute value of (2.4.5) is

$$O(|\theta|^{\beta+2-j}) \sum_{\nu=0}^{n-\beta-1} o(A_\nu^{\beta+1}) A_{n-\nu}^{\alpha-j} = O(|\theta|^{\beta+2-j}) o(n^{\alpha+\beta-j+2}) = o(n^\alpha),$$

and the absolute value of (2.4.6) is

$$O(\theta^2) \sum_{\nu=0}^{n-\beta-1} o(A_\nu^{\beta+1}) O(A_{n-\nu}^{\alpha-\beta}) = O(\theta^2) o(n^{\alpha+2}) = o(n^\alpha).$$

Hence $\theta P_n = o(n^\alpha)$, which in view of (2.4.1), (2.4.2) and (2.4.4) gives (2.3.7).

2.5. In order to prove the necessity of the condition in Lemma B, we have to show that, if the series (2.3.1) is summable (C^*, α) at the point $\theta = 0$, then the expression

$$\sum_{\nu=0}^{n} A_{n-\nu}^{\alpha} \nu b_\nu = \frac{1}{\theta} \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha} b_\nu \sin \nu\theta \cdot \frac{\nu\theta}{\sin \nu\theta} \tag{2.5.1}$$

is $o(n^{\alpha+1})$. We shall prove a slightly more general result, which will be required later on and which may be stated in the form of the following

LEMMA C. *Let $\{h_n\}$ be a sequence of numbers tending to 0 and satisfying for $n > n_0$ the condition*

$$0 < \delta \leqq n h_n \leqq \pi - \delta, \tag{2.5.2}$$

where δ is any fixed positive number less than $\frac{1}{2}\pi$. If for some such sequence the Cesàro means $\sigma_n^\alpha(\theta)$, where $\alpha > -1$, of the series (2.3.1) satisfy the condition

$$\sigma_\nu^\alpha(h_n) \to 0 \qquad \text{for } \nu \to \infty \text{ and } n \geqq \nu \tag{2.5.3}$$

then the sequence $\{\nu b_\nu\}$ is summable $(C, \alpha+1)$ to 0.

Proof. The argument does not differ appreciably from that of the preceding lemmas. We write

$$\begin{aligned} &\mu_\nu = \mu_\nu(\theta) = \frac{\nu\theta}{\sin \nu\theta}, \qquad \epsilon_\nu = \epsilon_\nu^{(n)} = A_{n-\nu}^{\alpha} \mu_\nu(h_n), \\ &v_n(\theta) = v_n^0(\theta) = \sum_{\nu=0}^{n} b_\nu \sin \nu\theta, \quad v_n^k(\theta) = \sum_{\nu=0}^{n} v_\nu^{k-1}(\theta), \qquad k = 1, 2, \cdots, \end{aligned} \tag{2.5.4}$$

and we suppose first that α is an integer. If we substitute $\theta = h_n$ into the right-hand side of (2.5.1) and denote the resulting expression by J_n, then (for $n > n_0$)

$$\begin{aligned} J_n &= h_n^{-1} \sum_{\nu=0}^{n} \epsilon_\nu b_\nu \sin \nu h_n \\ &= h_n^{-1} \sum_{\nu=0}^{n-\alpha-1} v_\nu^{\alpha}(h_n) \Delta^{\alpha+1} \epsilon_\nu + h_n^{-1} \sum_{k=0}^{\alpha} v_{n-k}^{k}(h_n) \Delta^k \epsilon_{n-k} = h_n^{-1} P_n + h_n^{-1} Q_n, \end{aligned} \tag{2.5.5}$$

say. From (2.5.3) we see that $v_\nu^\alpha(h_n) = o(\nu^\alpha) = o(n^\alpha)$. Hence

$$v_\nu^k(h_n) = o(n^\alpha) \qquad \text{for } \nu \to +\infty, \ n \geqq \nu, \ k = 0, 1, \cdots, \alpha,$$

and so (cf. (2.5.2))

$$h_n^{-1} Q_n = O(n) \sum_{k=0}^{\alpha} o(n^\alpha) \cdot O(1) = o(n^{\alpha+1}). \tag{2.5.6}$$

Since

$$\Delta^{\alpha+1}\epsilon_\nu = \sum_{j=0}^{\alpha+1} C_{\alpha+1,j} A_{n-\nu}^{\alpha-j} \Delta_\nu^{\alpha+1-j} \mu_{\nu+j}(h_n),$$

$h_n^{-1}P_n$ is a sum of expressions

$$O(|h_n|^{\alpha-j}) \sum_{\nu=0}^{n-\alpha-1} o(A_\nu^\alpha) A_{n-\nu}^{\alpha-j} = o(n^{\alpha+1}), \qquad j = 0, 1, \cdots, \alpha+1.$$

From this and from (2.5.5) and (2.5.6) we see that the left-hand side of (2.5.1) is $o(n^{\alpha+1})$, which proves Lemma C in the case of integral α.

2.6. If α is fractional, we write

$$\beta = [\alpha] + 1, \qquad \eta_\nu = \eta_\nu^{(n)} = A_{n-\nu}^{\alpha}(\mu_\nu(h_n) - \mu_n(h_n)),$$

where μ_ν has the same meaning as in (2.5.4). We get

$$J_n = h_n^{-1}\mu_n(h_n) \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha} b_\nu \sin \nu h_n + h_n^{-1} \sum_{\nu=0}^{n} \eta_\nu b_\nu \sin \nu h_n$$

$$= o(n^{\alpha+1}) + h_n^{-1} \sum_{\nu=0}^{n} \eta_\nu b_\nu \sin \nu h_n.$$

Furthermore,

$$h_n^{-1} \sum_{\nu=0}^{n} \eta_\nu b_\nu \sin \nu h_n = h_n^{-1} \sum_{\nu=0}^{n-\beta-1} v_\nu^\beta(h_n) \Delta^{\beta+1} \eta_\nu + h_n^{-1} \sum_{k=1}^{\beta} v_{n-k}^k(h_n) \Delta^k \eta_{n-k}$$

$$= h_n^{-1} P_n + h_n^{-1} Q_n.$$

From (2.5.3) and from the inequality $\beta > \alpha$ we get

$$v_\nu^\beta(h_n) = o(n^\beta) \qquad \text{for } \nu \to \infty,\ n \geqq \nu,$$

whence

$$v_\nu^k(h_n) = o(n^\beta) \qquad \text{for } \nu \to \infty,\ n \geqq \nu,\ k = 0, 1, \cdots, \beta.$$

Now

$$(2.6.1) \quad \Delta^k \eta_\nu = \sum_{j=0}^{k-1} C_{k,j} A_{n-\nu}^{\alpha-j} \Delta_\nu^{k-j} \mu_{\nu+j}(h_n) + A_{n-\nu}^{\alpha-k}(\mu_{\nu+k}(h_n) - \mu_n(h_n)),$$

so that

$$h_n^{-1} Q_n = h_n^{-1} \sum_{k=1}^{\beta} o(n^\beta) O(n^{-1}) = o(n^\beta) = o(n^{\alpha+1}).$$

If we substitute $k=\beta+1$ into (2.6.1), and observe that $h_n^{-1}P_n$ is correspondingly a sum of expressions

$$h_n^{-1}\sum_{\nu=0}^{n-\beta-1} o(A_n^{\beta})A_{n-\nu}^{\alpha-j}O(n^{j-\beta-1}) = o(n^{\alpha+1}), \qquad j = 0, 1, \cdots, \beta,$$

and of the expression

$$h_n^{-1}\sum_{\nu=0}^{n-\beta-1} o(A_n^{\beta})A_{n-\nu}^{\alpha-\beta-1}O((n-\nu)h_n) = o(n^{\alpha+1}),$$

we easily get

$$J_n = o(n^{\alpha+1}),$$

which completes the proof of Lemma B as well as that of Lemma C.

2.7. LEMMA D[11]. *Let $\sigma_n^{\alpha}(\theta)$ denote the Cesàro means of the series*

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty}(a_\nu \cos \nu\theta + b_\nu \sin \nu\theta). \tag{2.7.1}$$

A necessary and sufficient condition that this series should at the point θ_0 be summable (C^*, α), where $\alpha > -1$, to sum g, is that

(i) *this series should be summable (C, α) at the point θ_0 to sum g;*

(ii) *there should exist a sequence $\{h_n\}$ satisfying for $n > n_0$ the condition*

$$0 < \delta \leqq |nh_n| \leqq \pi - \delta, \qquad \delta \text{ independent of } n, \tag{2.7.2}$$

and such that

$$\sigma_\nu^{\alpha}(\theta_0 + h_n) \to g, \qquad \nu \to \infty, \ n \geqq \nu.$$

Proof. That the condition is necessary is obvious. In order to prove its sufficiency we may suppose that $\theta_0 = 0$. By hypothesis, the series

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} a_\nu$$

is summable (C, α) to g, and so on account of Lemma A the cosine part of the series (2.7.1) is summable (C^*, α) to g at the point $\theta = 0$. Hence, the sine part of (2.7.1) satisfies the hypothesis of Lemma C. It follows that the sequence $\{nb_n\}$ is summable $(C, \alpha+1)$ to 0. Consequently (Lemma B), the sine part of (2.7.1) is summable (C^*, α) to 0 at the point $\theta = 0$, and so the series (2.7.1) is summable (C^*, α) to g at that point. This completes the proof of Lemma D.

2.8. LEMMA E. *If the series (2.7.1) is summable (C^*, α), $\alpha > -1$, at a point θ_0, then the sequence*

(11) Cf. Footnote 8.

(2.8.1) $$nB_n(\theta_0) = n(b_n \cos n\theta_0 - a_n \sin n\theta_0)$$

is summable $(C, \alpha+1)$ *to* 0.

Assuming for simplicity that $\theta_0=0$, we see at once that Lemma E is a consequence of Lemmas A and B([12]).

3. Summability A^*. Lemmas on numerical series

3.1. We shall say that the series

(3.1.1) $$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta)$$

is summable A^* at a point θ_0 to sum g, if the harmonic function

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta)r^\nu$$

tends to g when the point $re^{i\theta}$ tends to $e^{i\theta_0}$ along any non-tangential path.

Lemma F. *If the series* (3.1.1) *is summable* (C^*, α), $\alpha > -1$, *at a point* θ_0 *to sum* g, *it is also summable* A^* *at that point to* g.

Proof. Let us assume that $\theta_0=0$. From the hypothesis of Lemma F it follows that both the series

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} a_\nu \cos \nu\theta, \qquad \sum_{\nu=0}^{\infty} b_\nu \sin \nu\theta$$

are summable (C^*, α) at the point θ, the sum of the first of these series being g, the sum of the second being 0. It is sufficient to show that

(3.1.2) $$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} a_\nu r^\nu \cos \nu\theta \to g, \qquad \sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta \to 0,$$

provided that

(3.1.3) $$z \to 1, \qquad |1 - z| \leqq C(1 - r), \qquad \text{where } z = re^{i\theta}.$$

Now,

(3.1.4) $$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} a_\nu r^\nu \cos \nu\theta = \Re\left(\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} a_\nu z^\nu\right),$$

([12]) From Lemmas E and G (see below) it follows at once that *if the series* (1.1.1) *is summable* (C^*, α), $\alpha > -1$, *at a point* θ_0, *and if the series* (1.1.2) *is summable A at that point, then the series* (1.1.2) *is also summable* (C^*, α) *at the point* θ_0.

This result, with summability A replaced by summability C, and with $\alpha=0, 1, \cdots$, is stated by Plessner [1].

and since the series $\frac{1}{2}a_0+a_1+a_2+\cdots$ is summable (C, α) to g, the very well known extension of the Abel-Stolz lemma gives(13) that, under the conditions (3.1.3), the function

$$\tfrac{1}{2}a_0 + a_1z + a_2z^2 + \cdots$$

tends to g. In view of (3.1.4), we get the first relation (3.1.2) under the same conditions.

Furthermore, denoting by u_ν^α the αth Cesàro means of the sequence $\{\nu b_\nu\}$, where $\nu=0, 1, \cdots$, we may write

$$\sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta = \int_0^\theta \left(\sum_{\nu=0}^{\infty} \nu b_\nu r^\nu \cos \nu t \right) dt = \int_0^\theta \Re\left(\sum_{\nu=0}^{\infty} \nu b_\nu \zeta^\nu \right) dt$$

$$= \int_0^\theta \Re\left\{ \left(\sum_{\nu=0}^{\infty} u_\nu^{\alpha+1} \zeta^\nu \right) (1-\zeta)^{\alpha+1} \right\} dt,$$

where $\zeta = re^{it}$. Taking into account that

$$|\zeta| = r, \qquad |1-\zeta| \leqq C(1-r), \qquad u_\nu^{\alpha+1} = o(\nu^{\alpha+1}),$$

we see that the absolute value of the last integral does not exceed

$$|\theta| \cdot o\left(\frac{1}{(1-r)^{\alpha+2}}\right) \cdot C^{\alpha+1}(1-r)^{\alpha+1} = o(1).$$

This proves the second relation (3.1.2) under the conditions (3.1.3). The proof of Lemma F is thus complete.

3.2. Lemma G which follows is known(14). We give its proof here for the sake of completeness only.

LEMMA G. *If the series*

$$\sum_{\nu=0}^{\infty} a_\nu \tag{3.2.1}$$

is summable A, and if the sequence $\{c_\nu\} = \{\nu a_\nu\}$ *is summable* $(C, \alpha+1)$ *to* 0, *where* $\alpha > -1$, *then the series* (3.2.1) *is summable* (C, α).

Proof. Let

$$\sigma_n^\gamma = \frac{s_n^\gamma}{A_n^\gamma}$$

denote the γth Cesàro means of the series (3.2.1). We shall say that the series (3.2.1) is *summable by the method* $A(C, \gamma)$ *to sum* s, if

(13) See §1.5 (ii).

(14) See, for example, Andersen [1].

$$\lim_{r\to 1-0}(1-r)\sum_{n=0}^{\infty}\sigma_n^{\gamma}r^n = s.$$

Summability $A(C, 0)$ is plainly identical with summability A. For the proof of Lemma G we shall require the following three supplementary propositions[15]:

(i) *If the series* (3.2.1) *is summable A to s, it is also summable* $A(C, \gamma)$ *to s, provided that* $\gamma \geqq 0$.

(ii) *If the series* (3.2.1) *is summable* $A(C, \gamma)$, *where* $\gamma > 0$, *to s, and if* $a_\nu = o(1/\nu)$, *then the series converges to s.*

(iii) *If* $a_\nu = o(1/\nu)$ *and if the series* (3.2.1) *is convergent, then it is summable* $(C, -1+\epsilon)$ *for every* $\epsilon > 0$.

In order to prove (i) we may assume that $\gamma > 0$ and that $s = 0$. We observe that

$$\frac{1}{A_n^{\gamma}} = \gamma\int_0^1 (1-u)^{\gamma-1}u^n du \qquad \gamma > 0;\ n = 0, 1, \cdots.$$

Hence, if

$$f(r) = \sum_{\nu=0}^{\infty} a_\nu r^\nu \qquad \text{for } 0 \leqq r < 1,$$

then

$$(1-r)\sum_{n=0}^{\infty}\sigma_n^{\gamma}r^n = (1-r)\sum_{n=0}^{\infty}\frac{s_n^{\gamma}}{A_n^{\gamma}}r^n$$

$$= (1-r)\gamma\int_0^1 (1-u)^{\gamma-1}\sum_{n=0}^{\infty}s_n^{\gamma}(ur)^n du \tag{3.2.2}$$

$$= (1-r)\gamma\int_0^1 \frac{f(ur)}{(1-ru)^{\gamma+1}}(1-u)^{\gamma-1}du.$$

Let M denote the upper bound of $f(x)$ for $0 \leqq x < 1$, and let us suppose that $|f(x)| \leqq \delta$ for $1-\epsilon \leqq x < 1$. The last expression in (3.2.2) does not exceed in absolute value

$$M(1-r)\gamma\int_0^{(1-\epsilon)r^{-1}}\frac{(1-u)^{\gamma-1}}{(1-ru)^{\gamma+1}}du + \delta(1-r)\gamma\int_{(1-\epsilon)r^{-1}}^1\frac{(1-u)^{\gamma-1}}{(1-ru)^{\gamma+1}}du.$$

Here the first term tends to 0 as $r\to 1$, and the second term does not exceed

$$\delta(1-r)\gamma\int_0^1\frac{(1-u)^{\gamma-1}}{(1-ru)^{\gamma+1}}du = \delta.$$

---

[15] Propositions (i) and (ii) are taken from Zygmund [1]. Proposition (iii) was proved by Hardy and Littlewood [1], even with o replaced by O.

(in order to verify the equation, it is sufficient to substitute $f(x)\equiv 1$, i.e., $\sigma_n^\gamma = 1$ for $n = 0, 1, \cdots$, into the extreme terms of (3.2.2)). This completes the proof of (i).

In order to prove (ii), it is sufficient to show that the summability $A(C, \gamma)$ of the series (3.2.1) and the condition $a_n = o(1/n)$ imply the summability A of (3.2.1), for then the result will follow from Tauber's classical theorem. We may of course assume that $\gamma \neq 0$. Let s_n denote the nth partial sum of the series (3.2.1). We shall show that then

$$\sigma_n^\gamma - s_n \to 0 \qquad \text{for } n \to \infty. \tag{3.2.3}$$

For

$$\sigma_n^\gamma - s_n = \frac{1}{A_n^\gamma} \sum_{\nu=0}^{n} (A_{n-\nu}^\gamma - A_n^\gamma) a_\nu, \tag{3.2.4}$$

and

$$A_n^\gamma - A_{n-\nu}^\gamma = \sum_{\mu=n-\nu+1}^{n} A_\mu^{\gamma-1}, \qquad 0 < \nu \leqq n. \tag{3.2.5}$$

Hence

$$0 \leqq A_n^\gamma - A_{n-\nu}^\gamma \leqq \nu A_n^{\gamma-1} \quad \text{for} \quad \gamma \geqq 1,$$

$$0 \leqq A_n^\gamma - A_{n-\nu}^\gamma \leqq \nu A_{n-\nu}^{\gamma-1} \quad \text{for} \quad 0 < \gamma \leqq 1.$$

Consequently, since $\nu a_\nu = o(1)$, we obtain

$$|\sigma_n^\gamma - s_n| \leqq \frac{A_n^{\gamma-1}}{A_n^\gamma} \sum_{\nu=1}^{n} o(1) = o(1), \qquad \text{if} \quad \gamma \geqq 1,$$

$$|\sigma_n^\gamma - s_n| \leqq \frac{1}{A_n^\gamma} \sum_{\nu=1}^{n} o(1) A_{n-\nu}^{\gamma-1} = o(1), \quad \text{if} \quad 0 < \gamma \leqq 1.$$

This completes the proof of (ii).

In order to prove (iii), we observe that

$$\sigma_n^{-1+\epsilon} = \frac{1}{A_n^{-1+\epsilon}} \sum_{\nu=0}^{[n/2]} A_{n-\nu}^{-1+\epsilon} a_\nu + \frac{1}{A_n^{-1+\epsilon}} \sum_{\nu=[n/2]+1}^{n} A_{n-\nu}^{-1+\epsilon} a_\nu = P_n + Q_n,$$

say. Now

$$|Q_n| \leqq \frac{1}{A_n^{-1+\epsilon}} \cdot o(1/n) \cdot \sum_{\nu=0}^{n} A_{n-\nu}^{-1+\epsilon} = o(1).$$

Denoting by s_n the partial sums of the series (3.2.1) and applying Abel's transformation we easily get

$$(3.2.6)\qquad P_n = \frac{1}{A_n^{-1+\epsilon}} \sum_{\nu=0}^{[n/2]-1} s_\nu A_{n-\nu}^{-2+\epsilon} + \frac{1}{A_{[n/2]}^{-1+\epsilon}} s_{[n/2]} A_{n-[n/2]}^{-1+\epsilon}.$$

Assuming, as we may, that $s_\nu \to 0$, we see that the last term on the right of (3.2.6) is $o(1)$. The first term on the right is

$$\frac{1}{A_n^{-1+\epsilon}} O(n^{-2+\epsilon}) \sum_{\nu=0}^{n} |s_\nu| = O(1/n) \sum_{\nu=0}^{n} |s_\nu| = o(1).$$

Hence

$$\sigma_n^{-1+\epsilon} \to 0,$$

which completes the proof of (iii).

3.3. We now pass on the proof of Lemma G. We may assume that the series is summable A to 0 and that $a_0 = 0$. Let $\sigma_n^\alpha = s_n^\alpha / A_n^\alpha$ denote the αth Cesàro means of the series (3.2.1), and let us suppose first that α is an integer. We write

$$va_\nu = c_\nu, \quad C_\nu^0 = c_\nu, \quad C_\nu^k = \sum_{\mu=0}^{\nu} C_\mu^{k-1}, \qquad \nu = 0, 1, 2, \cdots,$$

$$\epsilon_0 = 0, \qquad \epsilon_\nu = \epsilon_\nu^{(n)} = A_{n-\nu}^{\alpha} \frac{1}{\nu}, \qquad \nu = 1, 2, \cdots.$$

Then

$$(3.3.1)\qquad s_n^\alpha = \sum_{\nu=0}^{n} \epsilon_\nu c_\nu = \sum_{\nu=0}^{n-\alpha-1} C_\nu^{\alpha+1} \Delta^{\alpha+1} \epsilon_\nu + \sum_{k=0}^{\alpha} C_{n-k}^{k+1} \Delta^k \epsilon_{n-k} = P_n + Q_n,$$

say. From the formula

$$(3.3.2)\qquad \Delta^k \epsilon_\nu = \sum_{j=0}^{k} C_{k,j} A_{n-\nu}^{\alpha-j} \Delta^{k-j} \frac{1}{\nu+j}$$

and from the relation

$$C_n^{\alpha+1} = o(n^{\alpha+1})$$

it is easy to deduce that

$$(3.3.3)\qquad Q_n = o(n^\alpha).$$

Substituting $k = \alpha+1$ into (3.3.2) we get

$$\begin{aligned}\Delta^{\alpha+1} \epsilon_\nu &= A_{n-\nu}^{\alpha} \Delta^{\alpha+1} \frac{1}{\nu} + \sum_{j=1}^{k} C_{\alpha+1,j} A_{n-\nu}^{\alpha-j} \Delta^{\alpha+1-j} \frac{1}{\nu+j} \\ &= A_{n-\nu}^{\alpha} \Delta^{\alpha+1} \frac{1}{\nu} + \sum_{j=1}^{k} A_{n-\nu}^{\alpha-j} O\left(\frac{1}{\nu^{\alpha+2-j}}\right).\end{aligned}$$

Hence

$$P_n = \sum_{\nu=1}^{n-\alpha-1} A_{n-\nu}^{\alpha} C_\nu^{\alpha+1} \Delta^{\alpha+1} \frac{1}{\nu} + \sum_{j=1}^{\alpha+1} \sum_{\nu=1}^{n-\alpha-1} A_{n-\nu}^{\alpha-j} C_\nu^{\alpha+1} O\left(\frac{1}{\nu^{\alpha+2-j}}\right)$$

$$= \sum_{\nu=1}^{n-\alpha-1} A_{n-\nu}^{\alpha} C_\nu^{\alpha+1} \Delta^{\alpha+1} \frac{1}{\nu} + \sum_{j=1}^{\alpha+1} \sum_{\nu=1}^{n-\alpha-1} A_{n-\nu}^{\alpha-j} o(A_\nu^{j-1})$$

$$= \sum_{\nu=1}^{n-\alpha-1} A_{n-\nu}^{\alpha} C_\nu^{\alpha+1} \Delta^{\alpha+1} \frac{1}{\nu} + o(n^\alpha).$$

We may also write

$$P_n = \sum_{\nu=1}^{n} A_{n-\nu}^{\alpha} C_\nu^{\alpha+1} \Delta^{\alpha+1} \frac{1}{\nu} + o(n^\alpha).$$

From this and from (3.3.1) and (3.3.3) we get

$$\sigma_n^\alpha = \frac{1}{A_n^\alpha} \sum_{\nu=1}^{n} A_{n-\nu}^{\alpha} C_\nu^{\alpha+1} \Delta^{\alpha+1} \frac{1}{\nu} + o(1). \tag{3.3.4}$$

The first term on the right is the αth Cesàro mean τ_n^α of the series

$$0 + \sum_{\nu=1}^{\infty} C_\nu^{\alpha+1} \Delta^{\alpha+1} \frac{1}{\nu}, \tag{3.3.5}$$

whose terms are $o(1/\nu)$. By hypothesis, the series (3.2.1) is summable A to 0, and so on account of (i) summable $A(C, \alpha)$ to 0. From (3.3.4) it follows that the series (3.3.5) is summable $A(C, \alpha)$ to 0, and so converges to 0. Hence $\tau_n^\alpha \to 0$, which on account of (3.3.4) gives

$$\sigma_n^\alpha \to 0. \tag{3.3.6}$$

Lemma G is thus established in the case when α is an integer.

In the case of fractional $\alpha > -1$ the proof is similar. We write

$$\beta = [\alpha] + 1, \qquad \eta_0 = 0, \qquad \eta_\nu = \eta_\nu^{(n)} = A_\nu^\alpha \left(\frac{1}{\nu} - \frac{1}{n}\right) \qquad \text{for } 0 < \nu \leqq n.$$

Then

$$s_n^\alpha = \frac{1}{n} \sum_{\nu=1}^{n} c_\nu A_{n-\nu}^{\alpha} + \sum_{\nu=1}^{n} \eta_\nu c_\nu = o(n^\alpha) + \sum_{\nu=1}^{\beta} \eta_\nu c_\nu,$$

since the sequence $\{c_\nu\}$ is by hypothesis summable $(C, \alpha+1)$ to 0. On the other hand,

$$\sum_{\nu=1}^{n} \eta_\nu c_\nu = \sum_{\nu=1}^{n-\beta-1} C_\nu^{\beta+1} \Delta^{\beta+1} \eta_\nu + \sum_{k=0}^{\beta} C_{n-k}^{k+1} \Delta^k \eta_{n-k}.$$

Arguing as before (we omit the details, which remain essentially unchanged) we get instead of (3.3.4) the formula

$$\sigma_n^\alpha = \frac{1}{A_n^\alpha} \sum_{\nu=1}^{n} A_{n-\nu}^{\alpha} C_\nu^{\beta+1} \Delta^{\beta+1} \frac{1}{\nu} + o(1), \tag{3.3.7}$$

which in the case $\alpha \geqq 0$ gives (3.3.6) (the argument is then the same as in the case of integral α).

If $-1<\alpha<0$, then summability $(C, \alpha+1)$ of the sequence $\{c_\nu\}$ implies summability $(C, 1)$ of that sequence, to the limit 0, and so on account of the formula (3.3.4) with $\alpha=0$ we have

$$s_n = \sum_{\nu=1}^{n} C_\nu^1 \Delta \frac{1}{\nu} + o(1).$$

The series (3.2.1) being summable A, the series $\sum C_\nu^1 \Delta 1/\nu$, whose terms are $o(1/\nu)$, is convergent, and so summable (C, α) (cf. (iii)). In other words, the right-hand side of the formula (3.3.7), where $-1<\alpha<0$, $\beta=0$, tends to 0. This proves (3.3.6) also in the case $-1<\alpha<0$, and so Lemma G is proved.

3.4. LEMMA H. *If a series is summable A and its Cesàro means of order α are bounded $(\alpha>-1)$, then the series in summable $(C, \alpha+\epsilon)$ for any $\epsilon>0$.*

This lemma is well known(16).

LEMMA I. *If the Cesàro means $\sigma_n^\alpha=s_n^\alpha/A_n^\alpha$, where $\alpha>-1$, of a series $a_0+a_1+\cdots$ are bounded below, and if the series is summable A, then the series is summable $(C, \alpha+1)$.*

In the case $\alpha=0$ the lemma reduces to a very well known theorem of Littlewood. The general result may either be deduced from Littlewood's theorem by a certain comparatively simple argument(17), or may be proved exactly in the same way as the Littlewood theorem. We shall follow the latter course, using the familiar device of Karamata(18).

Without loss of generality we may assume that the expressions σ_n^α are all positive. By hypothesis,

$$\sum_{n=0}^{\infty} a_n r^n = (1-r)^{\alpha+1} \sum_{n=0}^{\infty} s_n^\alpha r^n \to s \qquad \text{as } r \to 1,$$

where s is the A-sum of the series $a_0+a_1+\cdots$. Replacing r by r^{k+1}, where k is any non-negative integer, we easily obtain from the last relation

(16) See Andersen [1]. A proof may also be found in Zygmund [1].
(17) See Zygmund [2, p. 329].
(18) Karamata [1].

$$(1-r)^{\alpha+1}\sum_{n=0}^{\infty} s_n^{\alpha} r^{(k+1)n} \to \frac{s}{(k+1)^{\alpha+1}} = \frac{s}{\Gamma(\alpha+1)}\int_0^1 x^k \left(\log\frac{1}{x}\right)^{\alpha} dx,$$

and so, if $P(x)$ is any polynomial,

$$(3.4.1)\quad (1-r)^{\alpha+1}\sum_{n=0}^{\infty} s_n^{\alpha} r^n P(r^n) \to \frac{s}{\Gamma(\alpha+1)}\int_0^1 P(x)\left(\log\frac{1}{x}\right)^{\alpha} dx.$$

Let $Q(x)$ be any function defined and bounded in the interval $0 \leqq x \leqq 1$. Let us assume furthermore that $Q(x)$ is continuous except at some point where it has a jump. Approximating $Q(x)$ above and below by polynomials and taking account of the positiveness of the expressions s_n^{α}, we deduce from (3.4.1) that

$$(3.4.2)\quad (1-r)^{\alpha+1}\sum_{n=0}^{\infty} s_n^{\alpha} r^n Q(r^n) \to \frac{s}{\Gamma(\alpha+1)}\int_0^1 Q(x)\left(\log\frac{1}{x}\right)^{\alpha} dx.$$

Let us define $Q(x)$ by the conditions

$$Q(x) = 0 \quad \textit{for} \quad 0 \leqq x < e^{-1}, \qquad Q(x) = \frac{1}{x} \quad \textit{for} \quad e^{-1} \leqq x \leqq 1.$$

The right-hand side of (3.4.2) is then equal to $s/\Gamma(\alpha+2)$. Hence, if we set $r = e^{-1/N}$, and make N tend to $+\infty$, the relation (3.4.2) gives

$$\frac{1}{N^{\alpha+1}}\sum_{n=0}^{N} s_n^{\alpha} \to \frac{s}{\Gamma(\alpha+2)},$$

or, since $A_n^{\alpha+1} \simeq n^{\alpha+1}/\Gamma(\alpha+2)$,

$$\frac{s_N^{\alpha+1}}{A_N^{\alpha+1}} \to s.$$

This completes the proof of the lemma.

3.5. Let $\sigma_n^{\alpha}(\theta)$ denote the Cesàro means of the trigonometric series

$$(3.5.1)\qquad \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\theta + b_n \sin n\theta).$$

The series will be said to be *finite* (C^*, α) at a point θ_0, if for any $A > 0$ there is a number $B = B(A)$, such that

$$|\sigma_n^{\alpha}(\theta_0 + h)| \leqq B \quad \textit{for} \quad |h| \leqq \frac{A}{n} \quad \textit{and} \quad n = 1, 2, \cdots.$$

Similarly the series (3.5.1) will be said to be *finite* A^* at the point θ_0, if for any $C > 0$ the harmonic function

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta) r^n$$

is bounded in the neighborhood of the point $e^{i\theta_0}$ satisfying the inequality $|re^{i\theta} - e^{i\theta_0}| \leqq C(1-r)$. It is plain that Lemmas D, E, F, G have the following analogues.

LEMMA D_1. *A necessary and sufficient condition that the series* (3.5.1) *should be finite* (C^*, α), $\alpha > -1$, *at the point* θ_0, *is that*

(i) *the series* (3.5.1) *should be finite* (C, α) *at the point* θ_0;

and that

(ii) *there should exist a sequence* $\{h_n\}$ *satisfying the condition*

$$0 < \delta \leqq |nh_n| \leqq \pi - \delta, \qquad \delta \text{ independent of } n,$$

and such that

$$\sigma_\nu^\alpha(\theta_0 + h_n) = O(1) \qquad \text{for } \nu \to \infty, \nu \leqq n.$$

LEMMA E_1. *If the series* (3.5.1) *is finite* (C^*, α), $\alpha > -1$, *at a point* θ_0, *then the sequence* $n(b_n \cos n\theta_0 - a_n \sin n\theta_0)$ *is finite* $(C, \alpha+1)$.

LEMMA F_1. *If the series* (3.5.1) *is finite* (C^*, α), $\alpha > -1$, *at the point* θ_0, *it is also finite* A^* *at that point.*

LEMMA G_1. *If the series* $a_0 + a_1 + \cdots$ *is finite* A, *and if the sequence* $\{va_\nu\}$ *is finite* $(C, \alpha+1)$, *then the series* $a_0 + a_1 + \cdots$ *is finite* (C, α) $(\alpha > -1)$.

4. PROOFS OF FUNDAMENTAL THEOREMS

4.1. LEMMA J. *If* G *is any measurable set of positive measure and* θ_0 *a point of density of* G, *then for any* $\lambda > 0$ *there is a sequence of numbers* $\{\alpha_n\}$ *such that*

$$\text{(4.1.1)} \qquad \begin{aligned} &n\alpha_n \to \lambda, \\ &\theta_0 \pm \alpha_n \quad \text{belongs to } G \text{ for all } n > n_0. \end{aligned}$$

Proof. Since θ_0 is a point of density, the average densities of the set G in the intervals

$$\left(\theta_0 + \frac{\lambda - \epsilon}{n}, \theta_0 + \frac{\lambda + \epsilon}{n}\right), \quad \left(\theta_0 - \frac{\lambda + \epsilon}{n}, \theta_0 - \frac{\lambda - \epsilon}{n}\right),$$

where $0 < \epsilon < \lambda$,

tend to 1 as n tends to infinity. It follows that for n large enough these average densities exceed $\frac{1}{2}$, and so there is a number β_n such that

$$\frac{\lambda - \epsilon}{n} \leqq \beta_n \leqq \frac{\lambda + \epsilon}{n}, \qquad \theta_0 \pm \beta_n \in G, \qquad n > n_0.$$

Thence it is not difficult to deduce the existence of a sequence $\{\alpha_n\}$ satisfying the conditions (4.1.1).

LEMMA K. *If the series*

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta) \tag{4.1.2}$$

is summable (C, α), $\alpha > -1$, *for* $\theta \in E$, *then the series is summable* (C^*, α) *almost everywhere in* E[19].

Similarly, if the series (4.1.2) *is finite* (C, α) *at every point of* E, *the series is finite* (C^*, α) *almost everywhere in* E.

Proof. Let us suppose that the series (4.1.2) is summable (C, α) for $\theta \in E$, and let $s(\theta)$ denote the (C, α)-sum of the series. Let G be any subset of positive measure of E, such that the Cesàro means $\sigma_n^{\alpha}(\theta)$ of the series (4.1.2) tend to $s(\theta)$ uniformly on G. In particular, $s(\theta)$ is continuous on G. Let $\theta_0 \in G$ be a point of density of G, and let $\{\alpha_n\}$ be any sequence of numbers such that

$$\tfrac{1}{4}\pi \leqq n\alpha_n \leqq \tfrac{3}{4}\pi, \qquad \theta_0 + \alpha_n \in G$$

for all $n > n_0$. Since

$$\sigma_{\nu}^{\alpha}(\theta_0 + \alpha_n) - s(\theta_0 + \alpha_n) \to 0 \qquad \text{for } \nu \to \infty,\ n \geqq \nu,$$

and since $s(\theta_0+\alpha_n) \to s(\theta_0)$, it follows that $\sigma_{\nu}^{\alpha}(\theta_0+\alpha_n) \to s(\theta_0)$ for $\nu \to \infty$, $n \geqq \nu$. An application of Lemma D shows that the series (4.1.2) is summable (C^*, α) at the point θ_0, that is, is summable (C^*, α) almost everywhere in G.

Since the measure of $E-G$ may be arbitrarily small, this proves the first part of the lemma. The second part may be proved in a similar way.

4.2. LEMMA L. *Let*

$$F(z) = u(r, \theta) + iv(r, \theta), \qquad z = re^{i\theta},$$

be a function regular inside the unit circle, and let us suppose that there is a set E *of positive measure situated on the circumference* $|z| = 1$ *and such that for any* $z_0 \in E$ *the function* $u(r, \theta)$ *is bounded when* z *tends to* z_0 *along non-tangential paths. Under these conditions for almost every point* $z_0 \in E$ *the functions* $u(r, \theta)$ *and* $v(r, \theta)$ *tend to finite limits when* z *tends to* z_0 *along any non-tangential paths.*

This lemma is known[20]. It plays an essential part in our proofs of Theorems 1 and 2. It may also be stated in the following form: *If a trigonometric series is finite* A^* *at every point of a set* E *of positive measure, then the series itself and its conjugate are summable* A^* *almost everywhere in* E.

[19] See Plessner [1] for $\alpha = 0, 1, 2, \cdots$.

[20] Privaloff [1]. A more general result, obtained by an argument similar to Privaloff's, will be found in Plessner [2].

4.3. From the preceding results we may deduce without difficulty the following special case of Theorem 1:

LEMMA M. *If a trigonometric series is summable* (C, β), $\beta > -1$, *at every point of a set* E, *the conjugate series is summable* (C, β) *almost everywhere in* E[21].

Proof. If the series (4.1.2) is summable (C, β) in E, then it is summable (C^*, β) almost everywhere in E (Lemma K). It follows (Lemma E) that

$$(4.3.1) \qquad (C, \beta+1)n(b_n \cos n\theta - a_n \sin n\theta) \to 0 \qquad \text{at almost every point of } E.$$

On the other hand, from Lemmas F and L it follows that the conjugate series

$$(4.3.2) \qquad \sum_{n=1}^{\infty} (a_n \sin n\theta - b_n \cos n\theta)$$

is summable A^*, and so also A, almost everywhere in E. From this and from (4.3.1) we see (cf. Lemma G) that the series (4.3.2) is summable (C, β) almost everywhere in E.

4.4. LEMMA N. *If at every point* θ *of a set* E *of positive measure the series*

$$(4.4.1) \qquad \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta)$$

is summable $(C, \alpha+1)$, *where* $\alpha > -1$, *to sum* $s(\theta)$, *and if*

$$(4.4.2) \qquad \sigma_\alpha(\theta) > -\infty \qquad \text{for } \theta \in E,$$

then

$$(4.4.3) \qquad \sigma^\alpha(\theta) < +\infty,$$

$$(4.4.4) \qquad s(\theta) = \tfrac{1}{2}\{\sigma_\alpha(\theta) + \sigma^\alpha(\theta)\}$$

at almost every point of E.

Proof[22]. Let us suppose first that α is an integer. From (4.4.2) and from Egoroff's well known theorem we deduce that there is a subset G of E, of measure differing as little as we please from that of E and such that

$$(4.4.5) \qquad \sigma_n^\alpha(\theta) > \sigma_\alpha(\theta) - \epsilon_n \qquad \text{for } \theta \in G,$$

where $\epsilon_1, \epsilon_2, \cdots$ is a sequence of numbers independent of θ and tending to 0. The functions $s(\theta)$ and $\sigma_\alpha(\theta)$ are measurable, and so if we remove from E subsets of arbitrarily small measure, these functions will be continuous on the

(21) See Plessner [1] (for the case $\alpha = 0, 1, \cdots$), Marcinkiewicz and Zygmund [1].

(22) Lemmas N and O of this paper are taken without essential changes from Marcinkiewicz and Zygmund [1].

remaining set. Without loss of generality we may assume that the functions $s(\theta)$ and $\sigma_\alpha(\theta)$ are continuous on G.

Let θ_0 be any point of density of the set G and belonging to G, and let $\{\beta_n\}$ be a sequence of numbers satisfying the conditions $n\beta_n \to \pi$, $\theta_0 \pm \beta_n \in G$ (cf. Lemma J). It follows from (4.4.5) that

$$\frac{1}{2}\{\sigma_n^\alpha(\theta_0+\beta_n)+\sigma_n^\alpha(\theta_0-\beta_n)\} > \frac{1}{2}\{\sigma_\alpha(\theta_0+\beta_n)+\sigma_\alpha(\theta_0-\beta_n)\}-\epsilon_n. \tag{4.4.6}$$

If we introduce the notation

$$C_0(\theta)=\tfrac{1}{2}a_0, \qquad C_\nu(\theta)=a_\nu\cos\nu\theta+b_\nu\sin\nu\theta \qquad \text{for } \nu \geqq 0,$$
$$\lambda_\nu=\lambda_\nu^{(n)}=A_{n-\nu}^\alpha\cos\nu\beta_n \qquad \text{for } 0 \leqq \nu \leqq n,$$

the left-hand side of the last inequality may be written

$$\begin{aligned}\frac{1}{2}\{\sigma_n^\alpha(\theta_0+\beta_n)+\sigma_n^\alpha(\theta_0-\beta_n)\} &= \frac{1}{A_n^\alpha}\sum_{\nu=0}^{n}A_{n-\nu}^\alpha C_\nu(\theta_0)\cos\nu\beta_n \\ &= \frac{1}{A_n^\alpha}\sum_{\nu=0}^{n}C_\nu(\theta_0)\lambda_\nu.\end{aligned} \tag{4.4.7}$$

Let $s_n^k(\theta)$ denote the kth Cesàro *sums* of the series (4.4.1). Applying summation by parts $\alpha+2$ times, we may represent the last sum in the form

$$\frac{1}{A_n^\alpha}\sum_{\nu=0}^{n-\alpha-2}s_\nu^{\alpha+1}(\theta_0)\Delta^{\alpha+2}\lambda_\nu+\frac{1}{A_n^\alpha}\sum_{j=0}^{\alpha+1}s_{n-j}^{j}(\theta_0)\Delta^j\lambda_{n-j}=P_n+Q_n,$$

say. Since $A_\mu^{-1}=A_\mu^{-2}=0$ for $\mu>0$, we may write

$$\Delta^{\alpha+2}\lambda_\nu=\sum_{j=0}^{\alpha+2}C_{\alpha+2,j}A_{n-\nu}^{\alpha-j}\Delta^{\alpha+2-j}\cos(\nu+j)\beta_n=\sum_{j=0}^{\alpha}C_{\alpha+2,j}A_{n-\nu}^{\alpha-j}\Delta^{\alpha+2-j}\cos(\nu+j)\beta_n$$
$$=\sum_{j=0}^{\alpha}A_{n-\nu}^{\alpha-j}O(\beta_n^{j-\alpha-2})=O(n^{-2})$$

for $0 \leqq \nu \leqq n-\alpha-2$. Assuming for simplicity that $s(\theta_0)=0$, that is, that

$$s_n^{\alpha+1}(\theta_0)=o(n^{\alpha+1}),$$

we see that P_n is equal to a sum of $\alpha+1$ expressions of the form

$$O\left(\frac{1}{n^2A_n^\alpha}\right)\sum_{\nu=0}^{n-\alpha-2}o(A_\nu^{\alpha+1})=o(1),$$

so that

$$P_n=o(1). \tag{4.4.8}$$

Since

$$\Delta^j \lambda_\nu = \sum_{h=0}^{j} C_{j,h} A_{n-\nu}^{\alpha-h} \Delta^{j-h} \cos(\nu + h)\beta_n \quad \text{for } 0 \leqq \nu \leqq n - j,$$

the coefficient of $s_{n-j}^j(\theta_0)$ in $A_n^\alpha Q_n$ is equal to

$$A_j^{\alpha-j} \cos n\beta_n + \sum_{h=0}^{j-1} C_{j,h} A_j^{\alpha-h} \Delta^{j-h} \cos(n - j + h)\beta_n$$

$$= A_j^{\alpha-j} \cos n\beta_n + \sum_{h=0}^{j-1} O(n^{h-j}) = A_j^{\alpha-j} \cos n\beta_n + O(n^{-1})$$

for $j = 0, 1, \cdots, \alpha+1$. On the other hand, the condition $s_m^{\alpha+1}(\theta_0) = o(m^{\alpha+1})$ implies

$$s_m^j(\theta_0) = o(m^{\alpha+1}) \qquad \text{for } j = 0, 1, \cdots, \alpha.$$

Hence,

$$Q_n = \frac{\cos n\beta_n}{A_n^\alpha} \sum_{j=0}^{\alpha+1} s_{n-j}^j(\theta_0) A_j^{\alpha-j} + o(1) = \frac{\cos n\beta_n}{A_n^\alpha} \sum_{j=0}^{\alpha} s_{n-j}^j(\theta_0) A_j^{\alpha-j} + o(1)$$

(since $A_j^{-1} = 0$ for $j > 0$).

The left-hand side of the inequality (4.4.6) is equal to $P_n + Q_n$, and so on account of (4.4.8) we get

$$\text{(4.4.9)} \qquad \tfrac{1}{2}\{\sigma_n^\alpha(\theta_0 + \beta_n) + \sigma_n^\alpha(\theta_0 - \beta_n)\} = \frac{\cos n\beta_n}{A_n^\alpha} \sum_{j=0}^{\alpha} s_{n-j}^j(\theta_0) A_j^{\alpha-j} + o(1).$$

The function $\sigma_\alpha(\theta)$ is by hypothesis continuous at the point θ_0 with respect to the set G. The right-hand side of the inequality (4.4.6) tends therefore to $\sigma_\alpha(\theta_0)$. From this and from (4.4.9) we deduce that

$$\liminf_{n\to\infty} \frac{\cos n\beta_n}{A_n^\alpha} \sum_{j=0}^{\alpha} s_{n-j}^j(\theta_0) A_j^{\alpha-j} \geqq \sigma_\alpha(\theta_0)$$

or, since $n\beta_n \to \pi$,

$$\limsup_{n\to\infty} \frac{1}{A_n^\alpha} \sum_{j=0}^{\alpha} s_{n-j}^j(\theta_0) A_j^{\alpha-j} \leqq -\sigma_\alpha(\theta_0).$$

4.5. Let us now consider the difference

$$\text{(4.5.1)} \qquad \sigma_n^\alpha(\theta_0) - \sigma_n^{\alpha+1}(\theta_0) = \sum_{\nu=0}^{n} C_\nu(\theta_0) \left\{ \frac{A_{n-\nu}^\alpha}{A_n^\alpha} - \frac{A_{n-\nu}^{\alpha+1}}{A_n^{\alpha+1}} \right\}$$

$$= \frac{1}{(n+\alpha+1)A_n^\alpha} \sum_{\nu=0}^{n} \nu A_{n-\nu}^\alpha C_\nu(\theta_0).$$

The last expression is analogous to (4.4.7), except that the factor $\cos \nu\beta_n$ is replaced by

$$\gamma_\nu = \gamma_\nu^{(n)} = \frac{\nu}{n+\alpha+1}, \tag{4.5.2}$$

which for our purposes possesses the same properties as $\cos \nu\beta_n$, namely

$$\Delta^k\gamma_\nu = O(n^{-k}), \qquad k = 0, 1, \cdots$$

(as a matter of fact, $\Delta^k\gamma_\nu = 0$ for $k > 1$). Hence, repeating the argument which gave (4.4.9), we get

$$\sigma_n^\alpha(\theta_0) - \sigma_n^{\alpha+1}(\theta_0) = \frac{n}{n+\alpha+1}\,\frac{1}{A_n^\alpha}\sum_{j=0}^{\alpha} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j} + o(1). \tag{4.5.3}$$

From (4.5.3) and from the inequality

$$\limsup_{n\to\infty} \frac{1}{A_n^\alpha}\sum_{j=0}^{\alpha} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j} \leqq -\sigma_\alpha(\theta_0),$$

we deduce that

$$\limsup_{n\to\infty} \left\{\sigma_n^\alpha(\theta_0) - \sigma_n^{\alpha+1}(\theta_0)\right\} \leqq -\sigma_\alpha(\theta_0),$$

or that

$$\sigma_\alpha(\theta_0) + \sigma^\alpha(\theta_0) \leqq 0.$$

This inequality was established under the hypothesis $\sigma_n^{\alpha+1}(\theta_0) \to 0$. In the general case, which may be reduced to the former by subtracting the constant $s(\theta_0)$ from the series (4.4.1), we obtain

$$\sigma_\alpha(\theta_0) + \sigma^\alpha(\theta_0) \leqq 2s(\theta_0). \tag{4.5.4}$$

This relation holds almost everywhere in G, and so almost everywhere in E. Hence (4.4.3) is true at almost every point of E.

Let us now change the signs of all the coefficients of the series (4.4.1) and let us apply the inequality (4.5.1) to the new series. We get that

$$\sigma_\alpha(\theta_0) + \sigma^\alpha(\theta_0) \geqq 2s(\theta_0) \tag{4.5.5}$$

almost everywhere in E, which on account of (4.5.1) gives (4.4.4) at almost every point of E. Thus Lemma N is proved in the case of integral α.

4.6. Passing to the case of fractional $\alpha > -1$, we set

$$\beta = [\alpha] + 1,$$

so that $\alpha < \beta < \alpha + 1$. Let us assume that the series (4.4.1) is summable not only $(C, \alpha+1)$, but also (C, β) on E. Let us define the set G in the same way as before, and let $\theta_0 \in G$ be any point of density of G. Let us assume for simplicity that

$$\sigma_n^\beta(\theta_0) \to 0.$$

Applying summation by parts $\beta+1$ times to the right-hand side of the equation

$$(4.6.1)\qquad \tfrac{1}{2}\{\sigma_n^\alpha(\theta_0+\beta_n)+\sigma_n^\alpha(\theta_0-\beta_n)\} = \frac{1}{A_n^\alpha}\sum_{\nu=0}^{n} C_\nu(\theta_0)\lambda_\nu$$

(cf. (4.4.7)) where $n\beta_n\to\pi$, $\theta_0\pm\beta_n\in G$, we get the expression

$$\frac{1}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)\Delta^{\beta+1}\lambda_\nu + \frac{1}{A_n^\alpha}\sum_{j=0}^{\beta} s_{n-j}^j(\theta_0)\Delta^j\lambda_{n-j} = P_n + Q_n,$$

say. To estimate P_n we use the formula

$$\Delta^{\beta+1}\lambda_\nu = \sum_{j=0}^{\beta} C_{\beta+1,j}A_{n-\nu}^{\alpha-j}\Delta^{\beta+1-j}\cos(\nu+j)\beta_n + A_{n-\nu}^{\alpha-\beta-1}\cos(\nu+\beta+1)\beta_n$$
$$\text{for } 0\leqq\nu\leqq n-\beta-1.$$

Correspondingly,

$$P_n = \sum_{j=0}^{\beta}\frac{O(n^{j-\beta-1})}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} o(A_n^\beta)A_{n-\nu}^{\alpha-j} + \frac{1}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)A_{n-\nu}^{\alpha-\beta-1}\cos(\nu+\beta+1)\beta_n$$

$$= o(1) + \frac{1}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)A_{n-\nu}^{\alpha-\beta-1}\cos(\nu+\beta+1)\beta_n.$$

On the other hand, the last expression is equal to

$$\frac{\cos n\beta_n}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \frac{1}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} o(A_\nu^\beta)\,|\,A_{n-\nu}^{\alpha-\beta-1}\,|\,O(n-\nu)\beta_n$$

$$= \frac{\cos n\beta_n}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \frac{\beta_n}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} o(A_\nu^\beta)O(A_{n-\nu}^{\alpha-\beta})$$

$$= \frac{\cos n\beta_n}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + o(1),$$

so that

$$P_n = \frac{\cos n\beta_n}{A_n^\alpha}\sum_{\nu=0}^{n-\beta-1} s_\nu^\beta(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + o(1).$$

Furthermore, from the formula

$$\Delta^j\lambda_\nu = \sum_{h=0}^{j} C_{j,h}A_{n-\nu}^{\alpha-j}\Delta^{j-h}\cos(\nu+h)\beta_n$$

we get

$$Q_n = \frac{1}{A_n^\alpha}\sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)\sum_{h=0}^{j-1} C_{j,h}A_j^{\alpha-j}O(\beta_n^{j-h}) + \frac{1}{A_n^\alpha}\sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j}\cos n\beta_n$$

$$= \frac{1}{A_n^\alpha}\sum_{j=0}^{\beta} o(n^\beta)O(n^{-1}) + \frac{\cos n\beta_n}{A_n^\alpha}\sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j}$$

$$= o(1) + \frac{\cos n\beta_n}{A_n^\alpha}\sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j}.$$

Hence the expression (4.6.1) is equal to

$$\frac{\cos n\beta_n}{A_n^\alpha}\left\{\sum_{\nu=0}^{n-\beta-1} s_\nu^{\beta}(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j}\right\} + o(1). \tag{4.6.2}$$

Since $n\beta_n \to \pi$, $\theta_0 \pm \beta_n \in G$, it follows that

$$\limsup_{n\to\infty} \frac{1}{A_n^\alpha}\left\{\sum_{\nu=0}^{n-\beta-1} s_\nu^{\beta}(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j}\right\} \leqq -\sigma_\alpha(\theta_0). \tag{4.6.3}$$

Let us now consider the equation (4.5.1). The argument which transformed (4.6.1) into (4.6.2) gives

$$\sigma_n^\alpha(\theta_0) - \sigma_n^{\alpha+1}(\theta_0)$$

$$= \frac{n}{n+\alpha+1}\,\frac{1}{A_n^\alpha}\left\{\sum_{\nu=0}^{n-\beta-1} s_\nu^{\beta}(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \sum_{j=0}^{\beta} s_{n-j}^{j}(\theta_0)A_j^{\alpha-j}\right\} + o(1). \tag{4.6.4}$$

Comparing this with (4.6.3) we get

$$\limsup_{n\to\infty}\left\{\sigma_n^\alpha(\theta_0) - \sigma_n^{\alpha+1}(\theta_0)\right\} \leqq -\sigma_\alpha(\theta_0),$$

or

$$\sigma^\alpha(\theta_0) + \sigma_\alpha(\theta_0) \leqq 0.$$

If $s(\theta_0) \neq 0$, this inequality is to be replaced by

$$\sigma^\alpha(\theta_0) + \sigma_\alpha(\theta_0) \leqq 2s(\theta_0). \tag{4.6.5}$$

This gives $\sigma^\alpha(\theta_0) < +\infty$, and so (4.4.3) is true almost everywhere in E. Hence applying (4.6.5) (with θ_0 replaced by general θ) to the series obtained from (4.4.1) by changing the signs of the coefficients of the latter series, we obtain

$$\sigma^\alpha(\theta) + \sigma_\alpha(\theta) \geqq 2s(\theta)$$

and so (4.4.4) is true almost everywhere in E.

4.7. We have therefore proved Lemma N also in the case of fractional α, but under an additional assumption, namely that the series (4.4.1) is sum-

mable not only $(C, \alpha+1)$, but also (C, β) in E. To remove this assumption we may argue as follows.

First of all, the argument of the preceding section shows that if

$$\sigma_\alpha(\theta) > -\infty, \qquad \sigma_n^\beta(\theta) = O(1) \tag{4.7.1}$$

for $\theta \in E$, then $\sigma^\alpha(\theta) < +\infty$, that is,

$$\sigma_n^\alpha(\theta) = O(1) \tag{4.7.2}$$

almost everywhere in E. On the other hand, since $\beta > \alpha$, the hypotheses of Lemma N imply that

$$\sigma_\beta(\theta) > -\infty, \qquad \sigma_n^{\beta+1}(\theta) \to s(\theta) \text{ on } E,$$

and applying Lemma N in the case of integral α (now β) we get that

$$\sigma_n^\beta(\theta) = O(1) \qquad \text{almost everywhere in } E.$$

It follows that under the conditions of Lemma N (in the case of fractional α) the relations (4.7.1) are true at almost every point of E, so that also (4.7.2) holds almost everywhere in E. Now, on account of Lemma H, the series (4.4.1) is summable (C, β) at every point θ at which

$$\sigma_n^\alpha(\theta) = O(1), \qquad \sigma_n^{\alpha+1}(\theta) \to s(\theta).$$

Hence, under the conditions of Lemma N, the series (4.4.1) is summable (C, β) almost everywhere in E. This completes the proof of Lemma N in the case of fractional α.

4.8. LEMMA O. *Suppose that for some $\alpha > -1$ the series*

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta) \tag{4.8.1}$$

satisfies at every point of a set E of positive measure the condition

$$-\infty < \sigma_\alpha(\theta) \leqq \sigma^\alpha(\theta) < +\infty, \tag{4.8.2}$$

and that the conjugate series

$$\sum_{\nu=1}^{\infty} (a_\nu \sin \nu\theta - b_\nu \cos \nu\theta) \tag{4.8.3}$$

is summable $(C, \alpha+1)$ in E. Then at almost every point of E the following conditions are satisfied:

$$-\infty < \tilde{\sigma}_\alpha(\theta) \leqq \tilde{\sigma}^\alpha(\theta) < +\infty, \tag{4.8.4}$$

$$\bar{s}(\theta) = \tfrac{1}{2}\{\tilde{\sigma}_\alpha(\theta) + \tilde{\sigma}^\alpha(\theta)\}, \tag{4.8.5}$$

$$\tilde{\sigma}^{\alpha}(\theta) - \tilde{\sigma}_{\alpha}(\theta) = \sigma^{\alpha}(\theta) - \sigma_{\alpha}(\theta). \tag{4.8.6}$$

Proof. The argument is so similar to that of Lemma N that we may condense some parts. Let G be a subset of E such that the functions $\sigma_{\alpha}(\theta)$ and $\sigma^{\alpha}(\theta)$ are continuous on G, and that

$$\sigma_{\alpha}(\theta) - \epsilon_n \leqq \sigma_n^{\alpha}(\theta) \leqq \bar{\sigma}_n^{\alpha}(\theta) + \epsilon_n \qquad \text{on } G,$$

where $\{\epsilon_n\}$ is a sequence of numbers tending to 0 and independent of θ. The measure of G may be as near to that of E as we please. Let θ_0 be any point of density of G belonging to G, and let β_n be a sequence of numbers satisfying the conditions

$$n\beta_n \to \tfrac{1}{2}\pi \qquad \theta_0 \pm \beta_n \in G. \tag{4.8.7}$$

Using the notation

$$\tilde{C}_0(\theta) = 0, \qquad \tilde{C}_{\nu}(\theta) = a_{\nu} \sin \nu\theta - b_{\nu} \cos \nu\theta \qquad \text{for } \nu > 0$$

we have the formula

$$\tfrac{1}{2}\{\sigma_n^{\alpha}(\theta_0 + \beta_n) - \sigma_n^{\alpha}(\theta_0 - \beta_n)\} = -\frac{1}{A_n^{\alpha}} \sum_{\nu=0}^{n} \tilde{C}_{\nu}(\theta_0) A_{n-\nu}^{\alpha} \sin \nu\beta_n. \tag{4.8.8}$$

Here the right-hand side is analogous to the expression (4.4.7), with $C_{\nu}(\theta_0)$ replaced by $\tilde{C}_{\nu}(\theta_0)$, and $\cos \nu\beta_n$ by $-\sin \nu\beta_n$. Let $\tilde{s}_{\nu}^{\prime}(\theta)$ denote the Cesàro sums of the series (4.8.3), and let us assume first that $\tilde{s}(\theta_0) = 0$.

In the case of integral α we get from (4.8.8) the formula

$$\tfrac{1}{2}\{\sigma_n^{\alpha}(\theta_0 + \beta_n) - \sigma_n^{\alpha}(\theta_0 - \beta_n)\} = -\frac{\sin n\beta_n}{A_n^{\alpha}} \sum_{j=0}^{\alpha} \tilde{s}_{n-j}^{j}(\theta_0) A_j^{\alpha-j} + o(1)$$

analogous to the formula for $P_n + Q_n$ in §4.4 (cf. (4.4.9)). Hence, making n tend to $+\infty$ and using (4.8.7) we obtain

$$-\liminf_{n\to\infty} \frac{1}{A_n^{\alpha}} \sum_{j=0}^{\alpha} \tilde{s}_{n-j}^{j}(\theta_0) A_j^{\alpha-j} \leqq \tfrac{1}{2}\{\sigma^{\alpha}(\theta_0) - \sigma_{\alpha}(\theta_0)\}. \tag{4.8.9}$$

Simultaneously we consider the formula for $\bar{\sigma}_n^{\alpha}(\theta_0) - \bar{\sigma}_n^{\alpha+1}(\theta_0)$, which is analogous to (4.5.1), with $C_{\nu}(\theta_0)$ replaced by $\tilde{C}_{\nu}(\theta_0)$. It follows that

$$\tilde{\sigma}_n^{\alpha}(\theta_0) - \tilde{\sigma}_n^{\alpha+1}(\theta_0) = \frac{n}{n+\alpha+1} \frac{1}{A_n^{\alpha}} \sum_{j=0}^{\alpha} \tilde{s}_{n-j}^{j}(\theta_0) A_j^{\alpha-j} + o(1)$$

(cf. (4.5.3)). From this and (4.8.9) we get

$$-\tilde{\sigma}_{\alpha}(\theta_0) \leqq \tfrac{1}{2}\{\sigma^{\alpha}(\theta_0) - \sigma_{\alpha}(\theta_0)\}.$$

If $\tilde{s}(\theta_0) \neq 0$, this formula takes the form

$$\bar{\sigma}_\alpha(\theta_0) - \bar{s}(\theta_0) \geqq -\tfrac{1}{2}\{\sigma^\alpha(\theta_0) - \sigma_\alpha(\theta_0)\}.$$

Applying this inequality, which is true for almost every point θ_0 of E, to the series (4.8.3) multiplied by -1, we get the inequality,

$$\bar{\sigma}^\alpha(\theta_0) - s(\theta_0) \leqq \tfrac{1}{2}\{\sigma^\alpha(\theta_0) - \sigma_\alpha(\theta_0)\}.$$

The last two inequalities show that the series (4.8.3) is finite at the point θ_0, and so almost everywhere in E.

Subtracting the last two inequalities we get the inequality

$$\bar{\sigma}^\alpha(\theta_0) - \bar{\sigma}_\alpha(\theta_0) \leqq \sigma^\alpha(\theta_0) - \sigma_\alpha(\theta_0),$$

true almost everywhere in E. The inequality opposite to this is also true, for the series (4.8.1) is summable $(C, \alpha+1)$ almost everywhere in E (Lemma M), so that under the hypotheses of Lemma O the series (4.8.1) and (4.8.3) play symmetric parts. (In particular, (4.8.5) is a consequence of the formula (4.4.4) in Lemma N.) It follows that (4.8.6) is true at the point θ_0, and so almost everywhere in E. This completes the proof of Lemma O in the case of integral α.

4.9. In the case of fractional $\alpha > -1$, the argument is identical with that of §4.6, provided we assume that the series (4.8.3) is not only summable $(C, \alpha+1)$, but summable (C, β), where $\beta = [\alpha]+1$, in E (or almost everywhere in E).

It is therefore sufficient to show that if α is fractional the hypotheses of Lemma O imply that the series (4.8.3) is summable (C, β) almost everywhere in E. First of all it may be remarked that if the series (4.8.1) is finite (C, α) and the series (4.8.3) finite (C, β) almost everywhere in E, then the series (4.8.3) is also finite (C, α) almost everywhere in E (the proof is substantially the same as the proof of the fact that if the series (4.8.1) is finite (C, α) and the series (4.8.3) summable (C, β) in E, then the latter series is finite (C, α) almost everywhere in E). On the other hand, the conditions of Lemma O imply that the series (4.8.1) is finite (C, β) and the series (4.8.3) summable $(C, \beta+1)$ in E, so that, using the lemma in the case of integral α, we see that the series (4.8.3) is finite (C, β), and so also finite (C, α), almost everywhere in E. Being finite (C, α) and summable $(C, \alpha+1)$ almost everywhere in E, the series (4.8.3) is summable (C, β) almost everywhere in E (Lemma M). This completes the proof of Lemma O in the case of fractional α.

4.10. Theorems 1 and 2 may now be proved in a few lines. Let us start with Theorem 1.

If the series (1.1.1) is summable A, and if $\sigma_\alpha(\theta) > -\infty$ at every point of a set E, then that series is summable $(C, \alpha+1)$ in E (Lemma I). The conjugate series (1.1.2) is then summable $(C, \alpha+1)$ almost everywhere in E (Lemma M). Theorem 1 is then a consequence of Lemmas N and O.

In order to prove Theorem 2, it is sufficient to show that under the hy-

potheses of that theorem the series (1.1.1) is summable A almost everywhere in E, for then Theorem 2 will follow from Theorem 1.

From the hypotheses of Theorem 2 it follows that the series (1.1.1) is finite (C^*, α) almost everywhere in E (Lemma K), and so finite A^* almost everywhere in E (Lemma F_1). From Lemma L it follows that the series (1.1.1) is summable A almost everywhere in E.

This completes the proofs of Theorems 1 and 2.

PART II

5. THE CASE OF POWER SERIES

5.1. By $K^0(\zeta; r)$ we shall denote the interior of the circle $K(\zeta; r)$, that is the set of points z such that $|z-\zeta| < r$. Similarly, by $A^0(\zeta; r, R)$ we shall denote the interior of the annulus $A(\zeta; r, R)$.

Let us suppose that the power series

$$\sum_{\nu=0}^{\infty} c_\nu e^{i\nu\theta} \tag{5.1.1}$$

is summable $(C, \alpha+1)$ at a point θ to sum $t(\theta)$. Let $\tau_n^\alpha(\theta)$ denote the αth Cesàro means of the series (5.1.1). It will be convenient to consider the set of the limit points of the sequence of numbers

$$\tau_n^\alpha(\theta) - t(\theta), \qquad n = 0, 1, 2, \cdots.$$

This set of limit points we shall denote by $L_1^\alpha(\theta)$. The set $L_1^\alpha(\theta)$ is obtained from $L_\alpha(\theta)$ by translating the latter by $-t(\theta)$.

The proof of Theorem 3 will be based on the following lemma.

LEMMA P. *Suppose that the series* (5.1.1) *is summable* $(C, \alpha+1)$ *(where* $\alpha > -1$*) for every* θ *of a set* E *of positive measure, and that for every* θ *of* E *the set*

$$L_1^\alpha(\theta)$$

does not contain any point of a fixed circle $K^0(\zeta; r)$, *with* $|\zeta| \geqq r$. *Then, for almost every* θ *of* E, *the set* $L_1^\alpha(\theta)$ *does not contain any point of the annulus* $A^0(0; |\zeta|-r, |\zeta|+r)$.

Proof. Let $\{\delta_n(\theta)\}$ denote the sequence of functions defined on the set E by the equation

$$\delta_n(\theta) = \max(0, \delta_n'(\theta)), \qquad \text{where } r - \delta_n'(\theta) = \inf_{\nu \geqq n} |\tau_\nu^\alpha(\theta) - t(\theta) - \zeta|\,(^{23}).$$

The functions $\delta_n(\theta)$ are measurable on E. They form a sequence decreasing monotonically to 0. For a fixed θ and for $\nu \geqq n$, the numbers $\tau_\nu^\alpha(\theta) - t(\theta)$ lie outside the circle $K^0(\zeta; r-\delta_n(\theta))$.

(23) By $\inf_{\nu \geqq n} c_\nu$ we mean the largest lower bound of the sequence $c_n, c_{n+1}, \cdots$.

Let G denote any subset of E on which the functions $\delta_n(\theta)$ tend uniformly to 0 and such that the function $t(\theta)$ is continuous on G. Let θ_0 denote any point of density of G belonging to G. Lemma P will have been proved when we have shown that the set $L_1^\alpha(\theta_0)$ contains no point of the annulus $A^0(0; |\zeta| - r, |\zeta| + r)$. Without loss of generality we may assume that

$$t(\theta_0) = 0. \tag{5.1.2}$$

Let λ be any real number and let $\{\beta_n\}$ denote a sequence of numbers such that

$$n\beta_n \to \lambda, \quad \theta_0 + \beta_n \in G \text{ for } n = 0, 1, 2, \cdots.$$

By $t_n^k(\theta)$ we shall denote the kth Cesàro sums of the series (5.1.1). On setting

$$C_\nu(\theta) = c_\nu e^{i\nu\theta} \qquad \text{for } \nu \geqq 0,$$

$$\mu_\nu = \mu_\nu^{(n)} = A_{n-\nu}^\alpha e^{i\nu\beta_n} \qquad \text{for } 0 \leqq \nu \leqq n,$$

we get the formula

$$\tau_n^\alpha(\theta_0 + \beta_n) = \frac{1}{A_n^\alpha} \sum_{\nu=0}^{n} A_{n-\nu}^\alpha C_\nu(\theta_0) e^{i\nu\beta_n} = \frac{1}{A_n^\alpha} \sum_{\nu=0}^{n} C_\nu(\theta_0)\mu_\nu.$$

This formula is entirely analogous to the formula (4.4.7). Since for our purposes the numbers $e^{i\nu\beta_n}$ have the same properties as the numbers $\cos \nu\beta_n$ (namely

$$\Delta^k e^{i\nu\beta_n} = O(|\beta_n|^k) \qquad \text{for } k = 0, 1, 2, \cdots),$$

the argument which led from (4.4.7) to (4.4.9) gives in the case of integral α the formula

$$\tau_n^\alpha(\theta_0 + \beta_n) = \frac{e^{in\beta_n}}{A_n^\alpha} \sum_{j=0}^{\alpha} t_{n-j}^j(\theta_0) A_j^{\alpha-j} + o(1). \tag{5.1.3}$$

Similarly, starting with the formula

$$\tau_n^\alpha(\theta_0) - \tau_n^{\alpha+1}(\theta_0) = \frac{1}{n+\alpha+1} \frac{1}{A_n^\alpha} \sum_{\nu=0}^{n} \nu A_{n-\nu}^\alpha C_\nu(\theta_0),$$

analogous to (4.5.1), we get the formula

$$\tau_n^\alpha(\theta_0) - \tau_n^{\alpha+1}(\theta_0) = \frac{n}{n+\alpha+1} \frac{1}{A_n^\alpha} \sum_{j=0}^{\alpha} t_{n-j}^j(\theta_0) A_j^{\alpha-j} + o(1) \tag{5.1.4}$$

corresponding to (4.5.3).

From (5.1.2), (5.1.3) and (5.1.4) it follows that

$$\tau_n^\alpha(\theta_0 + \beta_n) = e^{in\beta_n} \frac{n+\alpha+1}{n} \tau_n^\alpha(\theta_0) + o(1). \tag{5.1.5}$$

Since $\tau_n^{\alpha+1}(\theta_0) \to 0$ (cf. (5.1.2)), we have $t_n^{\alpha+1}(\theta_0) = o(n^{\alpha+1})$, so that

$$t_n^{\alpha}(\theta_0) = t_n^{\alpha+1}(\theta_0) - t_{n-1}^{\alpha+1}(\theta_0) = o(n^{\alpha+1}).$$

Hence

$$\tau_n^{\alpha}(\theta_0) = o(n).$$

It follows that the factor $(n+\alpha+1)/n$ on the right-hand side of (5.1.5) may be suppressed, and the formula may be rewritten in the form

$$e^{-in\beta_n}\tau_n^{\alpha}(\theta_0 + \beta_n) = \tau_n^{\alpha}(\theta_0) + o(1). \tag{5.1.6}$$

This formula was obtained under the condition (5.1.2). The general case may be reduced to this special one by subtracting $t(\theta_0)$ from the constant term of the series (5.1.1). The formula (5.1.6) then takes the form

$$e^{-in\beta_n}\{\tau_n^{\alpha}(\theta_0 + \beta_n) - t(\theta_0)\} = \{\tau_n^{\alpha}(\theta_0) - t(\theta_0)\} + o(1).$$

Since however the function $t(\theta)$ is by hypothesis continuous at the point θ_0 with respect to the set G, so that

$$t(\theta_0) - t(\theta_0 + \beta_n) \to 0,$$

we get

$$e^{-in\beta_n}\{\tau_n^{\alpha}(\theta_0 + \beta_n) - t(\theta_0 + \beta_n)\} = \{\tau_n^{\alpha}(\theta_0) - t(\theta_0)\} + o(1). \tag{5.1.7}$$

The formula (5.1.5), and so also the formula (5.1.7), holds in the case of fractional $\alpha > -1$, for then instead of (5.1.3) we have

$$\tau_n^{\alpha}(\theta_0 + \beta_n) = \frac{e^{in\beta_n}}{A_n^{\alpha}}\left\{\sum_{\nu=0}^{n-\beta-1} t_\nu^{\beta}(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \sum_{j=0}^{\beta} t_{n-j}^{j}(\theta_0)A_j^{\alpha-j}\right\} + o(1)$$

(cf. (4.6.2)), where $\beta = [\alpha]+1$, and instead of (5.1.4)

$$\tau_n^{\alpha}(\theta_0) - \tau_n^{\alpha+1}(\theta_0) = \frac{n}{n+\alpha+1}\frac{1}{A_n^{\alpha}}\left\{\sum_{\nu=0}^{n-\beta-1} t_\nu^{\beta}(\theta_0)A_{n-\nu}^{\alpha-\beta-1} + \sum_{j=0}^{\beta} t_{n-j}^{j}(\theta_0)A_j^{\alpha-j}\right\} + o(1)$$

(cf. (4.6.4)).

By hypothesis, the points $\theta_0+\beta_n$ belong to G, so that the number $\tau_n^{\alpha}(\theta_0+\beta_n) - t(\theta_0+\beta_n)$ lies outside the circle $K^0(\zeta; r-\delta_n)$, where

$$\delta_n = \sup \delta_n(\theta)\ (^{34}) \qquad \text{for } \theta \in G.$$

It follows that the left-hand side of the equation (5.1.7) does not belong to the circle $K^0(\zeta e^{-in\beta_n}; r-\delta_n)$. Since $n\beta_n \to \lambda$ and $\delta_n \to 0$, the set $L_1^{\alpha}(\theta_0)$ lies outside

(34) In other words, δ_n is the least upper bound of the sequence $\{\delta_n\}$ on G.

the circle $K^0(\zeta e^{-i\lambda}; r)$. It is now sufficient to observe that λ is an arbitrary real number, so that the set $L_1^\alpha(\theta_0)$ lies outside the annulus $A^0(0; |\zeta|-r, |\zeta|+r)$.

This completes the proof of Lemma P.

5.2. It is now easy to prove Theorem 3.

Let Z denote the set of points θ belonging to E and such that the corresponding set $L_1^\alpha(\theta)$ is not of circular structure. The sets $L_1^\alpha(\theta)$ are closed. For every θ of Z there is then a circle

$$K^0(\zeta; r) \qquad \text{with } |\zeta| \geqq r, \tag{5.2.1}$$

such that

(i) $L_1^\alpha(\theta)$ contains no point of $K^0(\zeta; r)$;

(ii) $L_1^\alpha(\theta)$ contains points of the annulus $A^0(0; |\zeta|-r, |\zeta|+r)$.

We may suppose that the circle (5.2.1) is rational, that is, the three numbers ξ, η, r, where $\xi+i\eta=\zeta$, are rational.

Let $Z_{\xi,r}$ denote the set of points θ of E satisfying conditions (i) and (ii). By Lemma P, every set $Z_{\xi,r}$ is of measure zero. Since

$$Z = \sum Z_{\xi,r},$$

where it is sufficient to extend summation over all rational circles $K^0(\zeta; r)$, the measure of Z is equal to 0.

This completes the proof of Theorem 3.

5.3. We now pass to the proof of Theorem 4. The theorem being obvious for $\alpha=0$, we may assume that $\alpha>0$. We have already observed in §1.4 that it is sufficient to prove the relation (1.4.2).

From the formula

$$A_n^\alpha = \frac{n+\alpha}{n} A_{n-1}^\alpha$$

we deduce that

$$\tau_n^\alpha(\theta) - \tau_{n-1}^\alpha(\theta) = c_n e^{in\theta} + \sum_{\nu=0}^{n} \left(\frac{A_{n-\nu}^\alpha}{A_n^\alpha} - \frac{A_{n-\nu-1}^\alpha}{A_{n-1}^\alpha} \right) c_\nu e^{i\nu\theta}$$

$$= c_n e^{in\theta} + \sum_{\nu=0}^{n-1} \frac{A_{n-\nu}^{\alpha-1}}{nA_n^\alpha} \nu c_\nu e^{i\nu\theta}.$$

It is therefore enough to show that the expression

$$\sum_{\nu=0}^{n-1} \frac{A_{n-\nu}^{\alpha-1}}{nA_n^\alpha} \nu |c_\nu| = \sum_{\nu=0}^{[n/2]} + \sum_{\nu=[n/2]+1}^{n} = S_1 + S_2,$$

say, tends to 0. Now

$$S_1 = \frac{O(n^{\alpha-1})}{nA_n^\alpha} \sum_{\nu=0}^{[n/2]} \nu |c_\nu| = O(n^{-2}) \sum_{\nu=0}^{n} \nu |c_\nu| = o(1),$$

$$S_2 = \frac{o(n)}{nA_n^\alpha} \sum_{\nu=[n/2]+1}^{n-1} A_{n-\nu}^{\alpha-1} = o(n^{-\alpha}) \sum_{\mu=0}^{n} A_\mu^{\alpha-1} = o(n^{-\alpha}) A_n^\alpha = o(1).$$

This completes the proof[25].

6. Strong summability of trigonometric series

6.1. We shall now prove a theorem on the strong summability of conjugate trigonometric series.

The series $\sum_{\nu=0}^{\infty} c_\nu$ is said to be *strongly summable* $(C, 1)$—or simply *strongly summable*—with index $q>0$, to sum s, if

$$\frac{1}{n+1} \sum_{\nu=0}^{n} |s_\nu - s|^q \to 0 \qquad \text{for } n \to \infty,$$

where s_ν denote the partial sums of the series considered.

Theorem 5. *If the series*

$$\tfrac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta) \tag{6.1.1}$$

is strongly summable, with index $q \geqq 1$, at every point of a set E, then the conjugate series

$$\sum_{\nu=1}^{\infty} (a_\nu \sin \nu\theta - b_\nu \cos \nu\theta) \tag{6.1.2}$$

is strongly summable, with index q, almost everywhere in E.

This theorem was stated by us in an earlier paper[26], with a sketch of proof. Since however the theorem has meanwhile been used in the proof of another result[27], we shall repeat here the proof of Theorem 7 in more detail.

Proof. Since $q \geqq 1$, the condition

$$\frac{1}{n+1} \sum_{\nu=0}^{n} |s_\nu(\theta) - s(\theta)|^q \to 0 \tag{6.1.3}$$

implies

$$\frac{1}{n+1} \sum_{\nu=0}^{n} |s_\nu(\theta) - s(\theta)| \to 0 \tag{6.1.4}$$

for $\theta \in E$. In particular, the series (6.1.1) is summable $(C, 1)$ in the set E. It

[25] A similar argument shows that the theorem holds in the case $-1<\alpha<0$, provided we replace the condition $c_n \to 0$ by $c_n = o(n^\alpha)$.

[26] See Marcinkiewicz and Zygmund [1].

[27] See Marcinkiewicz [1].

follows (Lemma M) that the series (6.1.2) is summable (C, 1) at every point of a set E_1 contained in E and of the same measure as E. Let $\bar{s}(\theta)$ denote the $(C, 1)$ sum of the series (6.1.2), and let G be a subset of E_1 such that

(i) the function $\bar{s}(\theta)$ is continuous on G,

(ii) the relation (6.1.3) holds uniformly on G.

Theorem 5 will have been proved when we have shown that (6.1.2) is strongly summable at every point θ_0 of G which belongs to G and is a point of density of G. Having fixed such a point θ_0, we may assume without loss of generality that

$$\tilde{\sigma}_n^1(\theta_0) \to 0.$$

Let $\{\alpha_n\}$ be any sequence of real numbers such that

$$n\alpha_n \to \tfrac{1}{2}\pi, \qquad \theta_0 \pm \alpha_n \in G \text{ for all } n > n_0. \tag{6.1.5}$$

From the formula

$$\begin{aligned} s_\nu(\theta_0 + \alpha_n) - s_\nu(\theta_0 - \alpha_n) &= -\sum_{\mu=1}^{\nu} (a_\mu \sin \mu\theta_0 - b_\mu \cos \mu\theta_0) \sin \mu\alpha_n \\ &= -\bar{s}_\nu(\theta_0) \sin \nu\alpha_n - \bar{s}_{\nu-1}^1(\theta_0)\Delta \sin (\nu-1)\alpha_n \\ &\quad - \sum_{\mu=1}^{\nu-2} \bar{s}_\mu^1(\theta_0)\Delta^2 \sin \mu\alpha_n \end{aligned}$$

it is easy to see that

$$s_\nu(\theta_0 + \alpha_n) - s_\nu(\theta_0 - \alpha_n) = -\bar{s}_\nu(\theta_0) \sin \nu\alpha_n + \epsilon_{\nu,n} \qquad \text{for } 0 \leqq \nu \leqq n,$$

where $\epsilon_{\nu,n}$ tends to 0 when $\nu \to \infty$, $n \geqq \nu$.

Hence

$$\begin{aligned} |\bar{s}_\nu(\theta_0)|^q |\sin \nu\alpha_n|^q \leqq \{ &|s_\nu(\theta_0 + \alpha_n) - s(\theta_0 + \alpha_n)| + |s(\theta_0 + \alpha_n) - s(\theta_0 - \alpha_n)| \\ &+ |s(\theta_0 - \alpha_n) - s_\nu(\theta_0 - \alpha_n)| + |\epsilon_{\nu,n}| \}^q \end{aligned}$$

and so

$$\begin{aligned} \frac{1}{n+1}\sum_{\nu=1}^{n} |\bar{s}_\nu(\theta_0)|^q |\sin \nu\alpha_n|^q \leqq{}& \frac{4^{q-1}}{n+1}\sum_{\nu=1}^{n} |s_\nu(\theta_0 + \alpha_n) - s(\theta_0 + \alpha_n)|^q \\ &+ \frac{4^{q-1}}{n+1}\sum_{\nu=1}^{n} |s(\theta_0 + \alpha_n) - s(\theta_0 - \alpha_n)|^q \\ &+ \frac{4^{q-1}}{n+1}\sum_{\nu=1}^{n} |s(\theta_0 - \alpha_n) - s_\nu(\theta_0 - \alpha_n)|^q \\ &+ \frac{4^{q-1}}{n+1}\sum_{\nu=1}^{n} |\epsilon_{\nu,n}|^q. \end{aligned} \tag{6.1.6}$$

The first and third terms on the right tend to 0 as $n\to\infty$, since the points $\theta_0\pm\alpha_n$ belong to the set G, on which the relation (6.1.3) holds uniformly. The second term on the right also tends to 0, for it does not exceed the expression

$$4^{q-1}\left|\, s(\theta_0+\alpha_n)-s(\theta_0-\alpha_n)\,\right|^q\to 0$$

(cf. condition (i)). Since the last term on the right obviously tends to 0, the left-hand side of (6.1.6) tends to 0. On account of the first relation (6.1.5) we get

$$\sum_{\nu=1}^{n}\left|\,\bar{s}_\nu(\theta_0)\,\right|^q\nu^q=o(n^{q+1}). \tag{6.1.7}$$

Let S_n denote the sum on the left of (6.1.7). Then

$$\sum_{\nu=1}^{n}\left|\,\bar{s}_\nu(\theta_0)\,\right|^q=\sum_{\nu=1}^{n}(S_\nu-S_{\nu-1})\nu^{-q}=\sum_{\nu=1}^{n-1}S_\nu(\nu^{-q}-(\nu+1)^{-q})+S_n n^{-q}$$

$$=\sum_{\nu=1}^{n-1}o(\nu^{q+1})O(\nu^{-q-1})+o(n^{q+1})n^{-q}=o(n).$$

This completes the proof of Theorem 5([28]).

6.2. A slight modification of the above argument gives the following more general result.

THEOREM 6. *Let $\phi(u)$, $u\geqq 0$, be a convex strictly increasing function vanishing for $u=0$. There is an absolute constant $\lambda>0$ with the following property. If, for any series* (6.1.1), *the sequence $\phi(\left|\,s_\nu(\theta)-s(\theta)\,\right|)$ is summable* $(C, 1)$ *to* 0 *at every point θ of a set E, the sequence $\phi(\lambda\left|\,\bar{s}_\nu(\theta)-\bar{s}(\theta)\,\right|)$, where $\bar{s}(\theta)$ is the* $(C, 1)$ *sum of* (6.1.2), *is summable* $(C, 1)$ *to* 0 *almost everywhere in E.*

We may take for example $\lambda=1/10$.

7. GENERALIZATION

7.1. Theorem 3 deals with power series; Theorems 1 and 2 with real parts of power series. It is natural to inquire how far the results we have obtained hold for Dirichlet's series

$$\sum_{\nu=1}^{\infty}c_\nu e^{-\lambda_\nu s},\qquad \lambda_1<\lambda_2<\cdots<\lambda_n\to\infty, \tag{7.1.1}$$

or, more generally for the Stieltjes integral of the type

$$\int_0^\infty e^{-\lambda s}dC(\lambda), \tag{7.1.2}$$

([28]) The above proof holds in the case $0<q<1$, if we add the condition that the series (6.1.2) is summable $(C, 1)$ in E.

where $C(\lambda)$ is a function which is of bounded variation over any finite interval, and $s=u+iv$ is a complex variable.

It is not difficult to show that the results we proved for power series hold mutatis mutandis for Dirichlet's series and for the more general integrals (7.1.2). The extensions must of course use the theory of arithmetic means of the integrals of the form (7.1.2).

The arithmetic means τ_w^α of order α for the integral (7.1.2) are defined by the formula

$$\tau_w^\alpha(s) = \int_0^w \left(1-\frac{\lambda}{w}\right)^\alpha e^{-\lambda s}\,dC(\lambda), \tag{7.1.3}$$

where w is a continuous variable tending to $+\infty$, and α is any non-negative number. In the case of a Dirichlet series the expression (7.1.3) reduces to M. Riesz's typical means.

Let $A(\lambda)$ and $B(\lambda)$ denote respectively the real and the imaginary part of the function $C(\lambda)$, so that

$$C(\lambda) = A(\lambda) - iB(\lambda).$$

If we suppose that the variable s is purely imaginary,

$$s = -iv,$$

then the real and imaginary parts of the integral (7.1.2) are respectively equal to the integrals

$$\int_0^\infty \cos \lambda v\,dA(\lambda) + \sin \lambda v\,dB(\lambda), \tag{7.1.4}$$

$$\int_0^\infty -\cos \lambda v\,dB(\lambda) + \sin \lambda v\,dA(\lambda), \tag{7.1.5}$$

and the second of these integrals will be spoken of as the *conjugate* to the first. If we denote the arithmetic means of order α for these two integrals by $\sigma_w^\alpha(v)$ and $\tilde{\sigma}_w^\alpha(v)$ respectively, then

$$\sigma_w^\alpha(v) = \int_0^w \left(1-\frac{\lambda}{w}\right)^\alpha (\cos \lambda v\,dA(\lambda) + \sin \lambda v\,dB(\lambda)),$$

$$\tilde{\sigma}_w^\alpha(\theta) = \int_0^w \left(1-\frac{\lambda}{w}\right)^\alpha (-\cos \lambda v\,dB(\lambda) + \sin \lambda v\,dA(\lambda)).$$

Let us write

$$\liminf_{w\to\infty} \sigma_w^\alpha(v) = \sigma_\alpha(v), \qquad \limsup_{w\to\infty} \sigma_w^\alpha(v) = \sigma^\alpha(v),$$

$$\sigma^\alpha(v) - \sigma_\alpha(v) = \omega_\alpha(v),$$

and let us define similarly the functions $\tilde{\sigma}_\alpha(v), \tilde{\sigma}^\alpha(v), \omega_\alpha(v)$ corresponding to the conjugate integral. Then the analogues of Theorems 1 and 2 for the integrals (7.1.4) may be stated as follows:

THEOREM 1′. *Suppose that the integral* (7.1.4) *is summable A to sum* $s(v)$, *at every point of a set* E *of positive measure, and that for an* $\alpha \geqq 0$

$$\sigma_\alpha(v) > -\infty, \qquad v \in E.$$

Then at almost every point of E *we have the relations*

$$\sigma^\alpha(v) < +\infty, \tag{7.1.6}$$

$$-\infty < \tilde{\sigma}_\alpha(v) \leqq \tilde{\sigma}^\alpha(v) < +\infty, \tag{7.1.7}$$

$$\tilde{\omega}_\alpha(v) = \omega_\alpha(v). \tag{7.1.8}$$

Moreover the integrals (7.1.4) *and* (7.1.5) *are summable by the arithmetic means of order* $>\alpha$ *at almost every point of the set* E *and*

$$s(v) = \tfrac{1}{2}\{\sigma_\alpha(v) + \sigma^\alpha(v)\}, \tag{7.1.9}$$

$$\tilde{s}(v) = \tfrac{1}{2}\{\tilde{\sigma}_\alpha(v) + \tilde{\sigma}^\alpha(v)\}, \tag{7.1.10}$$

almost everywhere in E, *where* $\tilde{s}(v)$ *denotes the value of the integral* (7.1.5) *by the method of arithmetic means of order* α.

THEOREM 2′. *If the arithmetic means* $\sigma_w^\alpha(v)$ *of order* α *of the integral* (7.1.4) *are bounded at every point of a set* E *of positive measure, then at almost every point of* E *the integrals* (7.1.4) *and* (7.1.5) *are summable by the arithmetic means of order* $>\alpha$, *and we have the relations* (7.1.7), (7.1.8), (7.1.9), (7.1.10).

In order to state the analogue of Theorem 3, let us suppose that the integral

$$\int_0^\infty e^{iv\lambda} dC(\lambda) \tag{7.1.11}$$

is at every point of a set E summable by the method of arithmetic means to the value $t(v)$. By $L^\alpha(v)$ we shall denote the set of limit points of the expression

$$\tau_w^\alpha(v) - t(v)\ (^{29})$$

(considered as a function of w), so that $L^\alpha(v)$ is the set of the numbers

$$\zeta = \lim_{n\to\infty} \{\tau_{w_n}^\alpha(v) - t(v)\}$$

(29) By $\tau_w^\alpha(v)$ we mean the expression

$$\int_0^w \left(1 - \frac{\lambda}{w}\right)^\alpha e^{iv\lambda} dC(\lambda).$$

where $\{w_n\}$ is an arbitrary sequence of real numbers tending to infinity. Then we have the following.

THEOREM 3′. *Suppose that at every point of a set E the integral* (7.1.11) *is summable by the method of arithmetic means of order* $\alpha+1$, *where* $\alpha \geqq 0$, *to the value* $t(v)$. *Then at almost every point of E the set* $L^{\alpha}(v)$ *is of circular structure with center* $t(v)$.

There is no need to give here the proofs of Theorems 1′, 2′, 3′, since they are wholly analogous to the proofs of Theorems 1, 2, 3. The theories of arithmetic means for ordinary series and for the integrals of the form (7.1.2) resemble so closely one another, and the technique of dealing with the arithmetic means of integrals is now so developed and so familiar (thanks mainly to the work of M. Riesz and G. H. Hardy([30])) that the passage from series to integrals usually requires no new ideas.

REFERENCES

A. F. ANDERSEN

1. *Studier over Cesàro's Summabilitetsmetode*, Copenhagen, Gjellerup, 1921.

G. H. HARDY AND J. E. LITTLEWOOD

1. *Contributions to the arithmetic theory of series*, Proceedings of the London Mathematical Society, vol. 11 (1913), pp. 401–478.

G. H. HARDY AND M. RIESZ

1. *Dirichlet's Series*, Cambridge Tracts, Cambridge, 1915.

J. KARAMATA

1. *Über die Hardy-Littlewoodsche Umkehrungen des Stetigkeitssatzes*, Mathematische Zeitschrift, vol. 32 (1930), pp. 319–320.

E. KOGBETLIANTZ

1. *Sommation des séries et intégrales divergentes*, Mémorial des Sciences Mathématiques, vol. 51, pp. 1–75.

B. KUTTNER

1. *A theorem on trigonometric series*, Journal of the London Mathematical Society, vol. 10 (1935), pp. 131–140.

J. MARCINKIEWICZ

1. *Sur la sommabilité forte des séries de Fourier*, Journal of the London Mathematical Society, vol. 14 (1939), pp. 162–168.

J. MARCINKIEWICZ AND A. ZYGMUND

1. *On the differentiability of functions and summability of trigonometrical series*, Fundamenta Mathematicae, vol. 26 (1936), pp. 1–43.

A. PLESSNER

1. *Über konjugierte trigonometrische Reihen*, Comptes Rendus de l'Académie des Sciences de l'URSS, vol. 4 (1935), pp. 251–253.
2. *Über das Verhalten analytischer Funktionen auf dem Rande des Definitionsbereiches*, Journal für die reine und angewandte Mathematik, vol. 158 (1928), pp. 219–227.

I. PRIVALOFF

1. *Sur une généralisation du théorème de Fatou*, Récueil Mathématique de Moscou, vol. 31 (1923), pp. 232–235.

([30]) See, for example, Hardy and Riesz [1].

W. ROGOSINKI

1. *Über die Abschnitte trigonometrischer Reihen*, Mathematische Annalen, vol. 95 (1925), pp. 110–134.

A. ZYGMUND

1. *Trigonometrical Series*, Monografie Matematyczne, vol. 5, Warsaw, 1935.
2. *Einige Sätze aus der Theorie der divergenten Reihen*, Bulletin de l'Académie Polonaise, 1928, pp. 309–331.

UNIVERSITY OF WILNO,
WILNO, POLAND
MT. HOLYOKE COLLEGE,
SOUTH HADLEY, MASS.

THE PARADOX OF KLEENE AND ROSSER[1]

BY

HASKELL B. CURRY

PART I. FORMULATION AND PRELIMINARY DISCUSSION

1. **Introduction.** In 1935 Kleene and Rosser published a proof that certain systems of formal logic are inconsistent, in the sense that every formula which can be expressed in their notation is also demonstrable[2]. There are, it so happens, only two systems in the literature to which this inconsistency proof applies; viz., the Church system of 1932–1933[3], and what I called an $\mathfrak{L}$-system in 1934[4]. But in spite of this limited application, the argument of Kleene and Rosser represents a theorem of great importance for the guidance of future research. It is a theorem of the same general character as the famous incompleteness theorems of Löwenheim, Skolem, and Gödel, which have played so prominent a role in recent research in mathematical foundations.

The proof of Kleene and Rosser is long and intricate[5], and contains complications which tend to obscure the essential meaning of their theorem. Accordingly there is some interest in the problem of making this paradox more accessible, and of presenting it in such a way as to make this essential meaning stand out more clearly. This is what the present paper attempts to do. The paradox is here presented and derived by a method which has many advantages, from the point of view indicated, over that of the original discoverers.

Before we enter into a detailed discussion, it will be expedient to examine the paradox in a vague preliminary way, and to explain in intuitive terms the central idea in its derivation.

One of the goals toward which mathematicians strive in setting up formal

Presented to the Society, March 27, 1937; received by the editors July 21, 1940. Minor additions were made December 7, 1940.

(1) This paper was first worked through in 1937. Since then it has been considerably revised. Part of this revision was done while I was in residence at the Institute for Advanced Study, Princeton, N. J. A general account of this subject was also presented to the Harvard Mathematical Colloquium on May 4, 1939. In the final preparation of the paper the author is greatly indebted to the referee, who checked the paper very carefully and conscientiously and made helpful suggestions.

(2) C 545.1. (For explanation of references, see end of the introduction.)

(3) C 359.4 and 6.

(4) C 396.6, p. 854.

(5) Although the paper C 545.1 is relatively short; yet the proof depends on results in a whole series of previous papers (viz., Church's C 359.4 and 6, Kleene's C 497.1 and 2, and Rosser's 546.1), totalling 162 pages. Of course some of the space in these papers is taken up with matters not considered here (cf. later parts of this introduction and the footnote to 3.15) and also with matters not bearing on the paradox.

systems is completeness—by which I mean not completeness in the technical sense, but simply the adequacy of the system for some purpose or other. There are two kinds of such completeness which especially concern us in this paper; both of them are desirable properties of formal systems of mathematical logic. I shall call these *combinatorial completeness* and *deductive completeness* respectively. They may be roughly explained as follows. A theory is combinatorially complete if and only if every expression $\mathfrak{M}$ formed from the terms of the system and an auxiliary indeterminate or variable x, can be represented within the system as a function of x (i.e., we can form in the system a function whose value for any argument is the same as the result of substituting that argument for x in $\mathfrak{M}$). A theory is deductively complete if whenever we can derive a proposition **B** on the hypothesis that another proposition **A** holds then we can derive without hypothesis a third proposition (such as $\mathbf{A} \supset \mathbf{B}$)(6) expressing this deducibility. Combinatorial completeness is thus a property relating to the possible constructions of terms (or formulas) within the system; deductive completeness relates to the possible derivations. Deductive completeness is a well known property of certain systems(7); whereas combinatorial completeness has only been achieved in recent years(8).

The essence of the Kleene-Rosser theorem is that it shows that these two kinds of completeness are incompatible—i.e., that any system which possesses both of them is inconsistent. The argument is essentially a refinement of the Richard paradox; it shows, in fact, that the Richard paradox can be set up formally within the system(9).

(6) Strictly speaking deductive completeness requires that when **A** and **B** are properties expressible by functions in the system such that **B** is formally derivable from **A**, then a formula expressing that fact, such as (x). $\mathbf{A}(x) \supset \mathbf{B}(x)$ is also derivable without hypothesis. However, the discussion in the text is sufficient for introductory purposes. For the precise formulation see 3.65 below.

(7) The name "deductive completeness" was suggested by the "Deduktionstheorem" of Hilbert-Bernays (C 507.1, p. 155), which establishes the deductive completeness of the ordinary logical calculus. The deductive completeness of the Church system is shown (essentially) by his Theorem I (C 359.4, p. 358); that of an $\mathfrak{L}$-system, by Theorem 14 of C 396.6.

(8) The possibility of achieving combinatorial completeness is due to Schönfinkel (C 304.1), who introduced certain operators, called "combinators" in C 396.2. An $\mathfrak{L}$-system, by definition, includes the theory of combinators, and so is combinatorially complete (in the strong sense—cf. below). The combinatorial completeness, according to the above rough definition, of the Church system was postulated in its formulation; that it can be so formulated as to be combinatorially complete (in the weak sense) according to the more precise definition of 3.3 (below) was shown by Rosser in C 546.1.

(9) Added in proof, August 21, 1941: Since this was written, I have discovered a much simpler way of deriving the contradiction for a system which is combinatorially complete in the strong sense (see 3.3 below). This new derivation which is based on the Russell paradox is contained in a paper, *The inconsistency of certain formal logics*, now in preparation. The new proof uses only the formulation of the present paper, viz. through 3.4.

In view of this circumstance the significance of most of the developments of this paper is quite different from what it seemed to be when this introduction was written. The uses which these developments have for other purposes is now where the chief interest lies. For these other

In order to see this in a preliminary way let us set up the Richard paradox([10]), as follows. In any formal system of arithmetic the number of definable numerical functions of natural numbers is enumerable; let them, therefore be enumerated, say in a sequence

$$\phi_1(x), \phi_2(x), \phi_3(x), \cdots . \tag{1}$$

Consider now the function

$$f(x) = \phi_x(x) + 1. \tag{2}$$

If this is a definable function, then it must be in the sequence (1), say it is ϕ_n; then putting ϕ_n for f and n for x in (2), we have

$$\phi_n(n) = \phi_n(n) + 1,$$

which is a contradiction.

Leaving aside the explanation of this paradox from an intuitive point of view, let us consider what happens in the case of a system which is both combinatorially and deductively complete. In such a system if a function is a numerical function—i.e., if it gives numerical values for all numerical arguments([11])—then a formal statement of that fact can be demonstrated within the system, since it is deductively complete; by means of a recursive enumeration of all the theorems the set of all numerical functions can then be effectively enumerated in a sequence (1). Since the theory is combinatorially complete, we can then define within the system the function f; this is demonstrably a numerical function, and the demonstration of this fact will tell us effectively the value of n such that the above contradiction will certainly arise.

This shows in a rough way the nature of the paradox. Before we proceed to the formal developments I shall interpolate a few remarks concerning the present proof and its relation to that of Kleene and Rosser.

---

purposes naturally certain modifications of the methods here given will eventually be necessary; what these modifications will be is for the future to determine. However, the following may be noted as significant now. First, wherever in this paper we have a theorem of the form

$$\vdash \mathcal{N}X_1 \& \cdots \& \mathcal{N}X_n \to \mathfrak{G}X_1 \cdots X_n$$

where $\mathfrak{G}$ does not contain F or $\mathcal{N}$, then on the basis of combinatory completeness alone it can be proved that

$$\vdash \mathfrak{G}\mathfrak{Z}_{p_1} \cdots \mathfrak{Z}_{p_n},$$

where $p_1, \cdots, p_n$ are any integers. Moreover where the second of these theorems is established here it follows from combinatorial completeness alone. This is true, in particular for 9.6; in fact under the hypothesis of 9.6 $T\mathfrak{Z}_n$ reduces to $\mathfrak{A}$. Second, we can go a certain, as yet undetermined, distance with a system, similar to that described in 5.6 for which consistency proof has now been obtained.

([10]) See also Church's discussion of this paradox in C 359.7.

([11]) More precisely, if from the assumption that the argument is numerical it follows formally that the functional value is numerical.

In the investigations of Church and his students the combinatorial completeness postulated is weakened in that it is required that the $\mathfrak{M}$ actually contain x—so that there need not be an apparatus for representing a constant as a function; there is also a weakening on the score of deductive completeness. These complications do not avoid the paradox, as Kleene and Rosser have shown; but they do increase considerably the length and intricacy of the derivation. If the object is to lay bare the central nerve of the paradox, the logical approach is to carry through the proof for the simpler case, that of combinatorial (and deductive) completeness in the strong sense, and then to show what modifications are necessary to carry through the proof for the more complicated case. Originally I had intended to do just this. But since the problems which led to this investigation, and the applications which I intend to make of it, are concerned with systems which are combinatorially complete in the strong sense, I have decided to leave this part of the investigation open. The present proof is, therefore, carried through in detail for the simpler case only. I believe that essentially the same method can be carried through with suitable modifications in the Church case, and that the result will have certain advantages over the original proof of Kleene and Rosser; but I leave the verification of this hunch to the readers who are more at home with such systems than I am.

Again, the present paper is essentially self-contained. I assume no acquaintance on the part of the reader with any previous papers either by Church or any of his students or myself. Reference is made to these papers for motivation or historical purposes only. The only information required of the reader which is not contained in the paper itself concerns the theory of recursive functions in ordinary arithmetic; in this case reference is made to papers by R. Péter, and such familiarity with the technique as is probably common knowledge among mathematicians is assumed. Except for these considerations regarding recursive functions (especially §6)([12]) all proofs are given in full.

The following system is used for the making of references. The paper is arranged according to the decimal system. Reference to different parts of the paper are made by these decimal numbers. References to other papers are made as follows: The letter "J" refers to the Journal of Symbolic Logic; it is followed by volume and page numbers where the paper referred to (or a review of it) may be found. The letter "C" refers to Church's *A bibliography of symbolic logic* in J 1, 121–218 and J 3, 178–212; the individual items are referred to according to the system explained in J 3, 193.

2. **Preliminary explanations.** In this section some general conventions regarding notation, and some considerations regarding formal systems in general, will be stated. Certain technical terms which are italicized are to be regarded as defined by the content in which they so appear.

([12]) On the proofs of these refer to the remarks at the beginning of 6.8.

2.1. *Formal systems.* The notion of formal system is the central concept in any formalistic view of mathematics. Such a system is characterized by the fact that the process of proof is specified by explicitly stated rules. The initial conventions specifying such a formal system I shall call collectively its *primitive frame.* This will consist of three kinds of conventions, as in 2.11–2.13.

2.11. First we shall have conventions stating what the objects of the theory, which I shall call its *terms*, shall be. These conventions will consist of a list of *primitive terms*, a list of *operations* for the formation of further terms and a set of *rules of formation* describing how new terms are to be formed from the primitive ones by use of the operations.

2.12. Next we have conventions specifying a set of propositions, which I shall call *elementary propositions*, concerning the terms. Ordinarily these conventions will consist simply of a list of *predicates*, representing categories of terms, relations between terms, etc.; and an elementary proposition is then formed by applying a predicate to the appropriate number of terms as arguments.

2.13. Finally we have specifications determining which of the elementary propositions are true. A certain set of these elementary propositions, called *axioms*, are stated to be true outright; and specific *rules of procedure* are given which are to determine how the derivation of new elementary theorems([13]) is to proceed. This process amounts to a recursive definition of the elementary theorems.

2.2. *Relations to symbolism.* This is not the place to go into an extended discussion of the nature of such a system; but it is necessary to make a few remarks, and, in particular, to compare it with the related notion of syntax of a language.

Any statement of the primitive frame must, of course, make use of symbols. However, these symbols are to be thought of as designating the term (or other notion)([14]) and not as a specimen of it. In other words, these symbols are not the objects of discourse, but refer to them; they are not constituents of some "language" whose "syntax" we are studying, but are technical terms to be adjoined to ordinary language as designating some objects or other having the properties set down for them in the primitive frame. The primitive frame does not specify what these objects are; nor is it necessary that it should do so. The essence of formal reasoning is that we use these symbols as nouns, verbs, etc., in our ordinary language without saying exactly what their meanings are.

Of course, the reader who wishes may interpret the theory as talking about symbols. In that case the symbols used in the primitive frame and our dis-

([13]) 'Theorem' is used here and elsewhere as synonymous with 'true proposition.'

([14]) The discussion here has primary reference to the symbols which are nouns. The modifications necessary for the other kinds of symbols are left to the reader.

cussions refer to these symbols being talked about. To avoid confusion it is desirable to use symbols of the second class distinct from those of the first. So interpreted the theory is essentially the syntax of an "object language." But this interpretation is not the only one possible, and the reader who insists on it must invent his own "object language." The symbols we are using become, in this interpretation, syntactical[15].

In order to have a systematic notation for names of symbols, Carnap (following Frege) uses a symbol enclosed in quotation marks as a name for that symbol (or expression). In this paper I shall, on occasion, use single quotes for this purpose, reserving double quotes for their ordinary uses (which sometimes conflict with the above technical use). But in some cases where there is no confusion, I shall use an expression to designate itself; in these cases the distinction is made clear by the context, as in the sentences 'Paris is a city' and 'Paris is dissyllabic'[16].

2.3. *Propositions.* 'Proposition' is here used in an intuitive sense, i.e., as designating the meaning of a sentence.

The elementary propositions are specified by 2.12; and 2.13 gives a recursive determination of all of them which are true. This, incidentally, defines the fundamental predicates uniquely; so that once we know what the terms are, no further indeterminateness in regard to the predicates exists.

Further propositions may be formed from the elementary propositions by using the connectives and processes of ordinary discourse. These will be called *metapropositions*, and the study of the theory by means of them metatheoretic[17]. Examples of them are given below.

---

[15] For a further discussion of the nature of a formal system, see my address *Some aspects of the problem of mathematical rigor*, Bulletin of the American Mathematical Society, vol. 47 (1941), pp. 221–241.

[16] This example is Carnap's (J 4,82—p. 153). Most of the cases of this "autonymous" mode of speech in the present paper will be found to be concerned with symbols (or expressions) which are not nouns—e.g. sentences or verbs (such as '$\vdash$'). When such symbols are used as nouns I omit the quotes systematically when the reference is to the meaning of the symbol, rather than its shape (or sound). Whether this is an instance of the "autonymous" mode of speech is a matter which I leave for the reader; cf. the preceding footnote.

[17] This is not the usual sense of 'metatheoretic.' In the ordinary usage, in fact, 'metatheoretic' and 'theoretic' are applied not to propositions but to expressions—i.e., rows of symbols; the distinction between them corresponds to the distinction between syntax language and object language. Cf. 2.2 and the paper cited in Footnote 15.

The connectives of ordinary discourse are, of course, to be understood as including processes which, as they occur in ordinary discourse, are only vaguely formulated. The precise definition of certain metatheoretic processes may require passing to a higher level of formalization. Examples of such processes are the deducibility connectives of 2.4 (cf. the next footnote) and substitution. (On the formulation of the latter, however, compare my paper cited in Footnote 15.) Consequently there is a certain amount of arbitrariness in the distinction between elementary and metatheoretic propositions. Whether we regard a proposition as metatheoretic for a system $\mathfrak{S}$, or as elementary for a more extended system $\mathfrak{S}'$, is a matter of the point of view. Although the latter may be more exact, yet often we prefer the former for the same reason that

2.4. *Metatheoretic connectives.* The following are used as symbols having a significance independent of the primitive frame of the system(18).

2.41. '$\equiv$' (for definitional identity).

2.42. '$\rightarrow$' (for deducibility; the proposition on the right follows from that on the left by the rules of procedure).

2.43. '&' (for intuitive conjunction).

2.44. '$\rightleftarrows$' (for equivalence).

2.45. Dots will be used for avoiding parentheses in the customary manner (cf. J 2, 26–28). The order of seniority of the connectives for this purpose is '$\equiv$,' '$\rightleftarrows$,' '$\rightarrow$,' '&,' '$=$' (defined in §3.2).

2.5. *Term extensions.* One method of forming metapropositions concerning a system $\mathfrak{S}$ is to consider the adjunction to it of indeterminates, precisely as in algebra we adjoin indeterminates to a field to form its transcendental extension. If 'X' is not defined with reference to a system $\mathfrak{S}$, then we can adjoin to the primitive frame for $\mathfrak{S}$ the specification 'X is a term.' If nothing else is specified concerning X we call such an extension a *term extension* of $\mathfrak{S}$ with respect to the indeterminate X, and denote it by $\mathfrak{S}(X)$. Similar notations will be used for extensions with respect to several indeterminates.

The following are some special conventions regarding indeterminates.

2.51. 'Term' will refer to any term of $\mathfrak{S}$ or its term extensions; a *constant*, however, is a term belonging to the underlying system.

2.52. By an *X-term* we shall mean a term of $\mathfrak{S}(X)$; similarly an $(X_1, \cdots, X_m)$-term is a term of $\mathfrak{S}(X_1, \cdots, X_m)$.

---

we do not rush to Paris when we wish to measure a meter. Thus, the elementary propositions of the system $\mathfrak{r}$ of §6 formalize what are ordinarily satisfactorily thought of as metatheorems of a more elementary system of arithmetic. It is even possible to conceive of a system in which there are no elementary propositions, all propositions being metatheoretic; certain constructionist theories amount, essentially, to just this.

(18) The exact definition of these metatheoretic connectives is discussed in my paper *Some properties of formal deducibility* (not yet published). For the purposes of the present paper it will, in general, be sufficient to take them as intuitively given. However, in regard to the connective $\rightarrow$ it seems to be necessary to make the following remarks.

As between elementary propositions we may distinguish two sorts of implication of a metatheoretic nature. Let **A** and **B** be two such propositions. Then it may happen that if **A** is adjoined to the primitive frame as additional axiom, then **B** is a theorem of the resulting system. This is the relation symbolized by **A**$\rightarrow$**B**; and it is generally what is referred to in connection with such phrases as '(metatheoretic) implication,' 'implies,' 'follows from,' etc. Moreover, it holds between **A** and **B** whenever **B** follows directly from **A** by a rule of procedure; hence, it is legitimate to state the rules of procedure in terms of $\rightarrow$. On the other hand **A**$\rightarrow$**B** does not exhaust the possibilities of implication relations between **A** and **B**; we may, for example, have a process for converting a proof of **A** into a proof of **B**. The most general implication relation is expressed here as "if **A**, then **B**;" this is to be regarded as including the possibility **A**$\rightarrow$**B** as a special case.

The difference between these two may be illustrated by an example. Consider for instance the classical algebra of propositions; and let p be a variable and $\mathfrak{A}$ an arbitrary formula therein. Then if the algebra is formulated with a substitution rule, we have $\vdash p \rightarrow \vdash \mathfrak{A}$, otherwise not. On the other hand the proposition

2.53. If a proposition holds when $X_1, \cdots, X_m$, which may appear in it, are indeterminate, that fact may be emphasized by attaching a subscript 'X' to the principal connective or predicate in it.

2.6. *Substitution.* If $\mathfrak{X}$ is an $(X_1, \cdots, X_m)$-term, and if $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ are terms, then the notation

$$\begin{bmatrix} \mathfrak{A}_1, \cdots, \mathfrak{A}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}$$

denotes the term $\mathfrak{Y}$ got by replacing X_i wherever it occurs in $\mathfrak{X}$ by $\mathfrak{A}_i$—i.e., the term which is constructed from $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ by the process that leads from $X_1, \cdots, X_m$ to $\mathfrak{X}$. We say that $\mathfrak{Y}$ is obtained from $\mathfrak{X}$ by substitution of $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ for $X_1, \cdots, X_m$. We evidently have the following property.

2.61. Any relation between the $(X_1, \cdots, X_m)$ terms $\mathfrak{X}_1, \mathfrak{X}_2, \cdots, \mathfrak{X}_n$ which is valid in $\mathfrak{S}(X_1, \cdots, X_m)$ will also hold, in the appropriate term extension, between the terms obtained by substituting any terms $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ for $X_1, \cdots, X_m$ throughout.

2.7. *Use of letters*(19).

---

if $\vdash \mathfrak{p}$ then $\vdash \mathfrak{A}$

is vacuous because the hypothesis is false; but we may note that it is a special case of a general substitution theorem which is demonstrable even when the algebra is formulated in terms of axiom schemes without a substitution rule.

This example shows that implications of the if-then type are ambiguous when the hypothesis is not verified. Usually we are only interested in such theorems when it is. In this paper theorems concerning $\mathfrak{S}$ stated in the if-then form are to be understood as making no particular assertion unless the hypotheses are verified; and theorems are occasionally stated in that form when that is all that interests us, even though the relation may actually be one of deducibility. Exceptions to this rule, made mostly for typographical reasons, are indicated by an asterisk prefixed to the number; such theorems are deducibility theorems. For the system $\mathfrak{r}$ all theorems and rules are stated in the if-then terminology.

In the paper above cited I have considered deducibility relations between propositions which are themselves compounded out of simpler ones by the above connectives and some others; and have advanced reasons for assuming these connectives subject to the rules of the Heyting calculus. If this is done, then I have shown that, for systems like $\mathfrak{S}$, the relation of deducibility holds between two elementary propositions if and only if the conclusion is deducible from the premise by the axioms and rules of procedure of the system. However, these general results are not needed here; we only need $\rightarrow$ between propositions which are conjunctions of elementary ones.

It should be noted that theorems in terms of deducibility which are true for $\mathfrak{S}$ are true for any extension of it.

The notations adopted above for distinguishing between deducibility and other forms of implication are not always rigidly adhered to. This is because when the paper was first written the distinction was left to be determined from the context. The changes made since will assist the reader at the more important points; but it is difficult to be sure that every instance of ambiguity has been caught in the revision.

(19) These conventions do not apply to the introduction, nor to certain intuitive remarks not strictly necessary for logical completeness.

2.71. Capital letters are used for notations connected with the formal system $\mathfrak{S}$ and its various term extensions, as follows:

2.711. 'U,' 'V,' 'X,' 'Y,' 'Z' are for indeterminates.

2.712. Other Latin, and also Greek or script, capitals are for specific constants. Their meaning is fixed throughout the paper.

2.713. '$\mathfrak{S}$' designates the system.

2.714. '$\mathfrak{F}$' and '$\mathfrak{T}$' designate specific functions from numerical expressions to terms. Their meaning is also fixed throughout the paper.

2.715. Other German letters[20] are for unspecified terms, or for terms which are specified only for a particular context. They are the variables of the intuitive discussion.

2.72. Lower case letters and Arabic numerals are for natural numbers and notions related to them, as follows:

2.721. Arabic numerals are for natural numbers in the ordinary sense.

2.722. 'u,' 'v,' 'x,' 'y,' 'z' are for numerical variables.

2.723. Other Latin letters are for unspecified numbers either taken intuitively or as constituents of the system of §6.

2.724. Greek letters[21] are for specific numerical functions. They have a fixed meaning throughout the paper.

2.725. $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, $\mathfrak{d}$ are for unspecified numerical expressions (cf. §6).

2.726. $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{k}$ are for unspecified functions.

2.727. $\mathfrak{m}$, $\mathfrak{p}$, $\mathfrak{q}$ are for unspecified expressions whether functional or numerical.

2.728. $\mathfrak{e}$, $\mathfrak{f}$, $\mathfrak{n}$, $\mathfrak{o}$, $\mathfrak{r}$, $\mathfrak{v}$, $\mathfrak{x}$, $\mathfrak{y}$, $\mathfrak{z}$ are for miscellaneous special purposes. The first six of these have a fixed meaning; the others are for unspecified classes of numerical variables.

2.729. Gothic letters are for operations of the system $\mathfrak{r}$ in §6.

3. **Formulation of the system $\mathfrak{S}$.** The assumptions concerning the system $\mathfrak{S}$ are those stated in 3.1–3.6. It will be noted that, aside from the very general assumptions in 3.1, the essential restrictions in the system are those of 3.3 and 3.6, which represent combinatorial and deductive completeness respectively. The other assumptions are essentially either definitions of particular terms of $\mathfrak{S}$ (and so are primarily notational) or consequences of the other assumptions.

3.1. *General restrictions on the primitive frame.*

3.11. There are a finite number of primitive terms, which we shall designate $E_0, E_1, \cdots, E_t$. (That this is not an essential restriction is shown in 3.123.)

3.12. The only operation is a single binary one of universal applicability.

---

[20] '$\mathfrak{L}$' is used in connection with 'system' to refer to the system so defined in C 396.6, p. 854 and some of its modifications.

[21] Except 'ϵ' which is used in the sense of Gödel (C 418.3, p. 180) in 6 and also (in the form '$\in$') in its ordinary set-theoretic meaning, and 'λ' which is given a specific meaning in 3.3.

We shall call this *application* and denote the term formed by application of $\mathfrak{A}$ to $\mathfrak{B}$ simply by $\mathfrak{AB}$. Then we have as sole rule of formation:

If $\mathfrak{A}$ and $\mathfrak{B}$ are terms, $\mathfrak{AB}$ is a term.

3.121. Parentheses will be used in the customary manner to indicate the construction of complex terms. To avoid superfluous parentheses, however, it shall be understood that in a row of terms without parentheses the operation shall be performed from left to right; so that $\mathfrak{ABCD}$, for instance, means the same as $((\mathfrak{AB})\mathfrak{C})\mathfrak{D}$.

3.122. The intuitive meaning of this operation is that of application of a function to an argument; thus if $\mathfrak{A}$ is a function $\mathfrak{g}$ and $\mathfrak{B}$ is a value b of its argument, $\mathfrak{AB}$ is what we ordinarily write as $\mathfrak{g}(b)$; if $\mathfrak{A}$ is a function of two variables, such as the addition function $x+y$, and $\mathfrak{B}$ is a value b, then $\mathfrak{AB}$ is the function of one variable got by giving the first argument in $\mathfrak{A}$ the value b—i.e., the function $b+x$—; this can be applied again to an argument $\mathfrak{C}$ representing a value c, and the result, $\mathfrak{ABC}$, is $b+c$, etc. This notion, due to Schönfinkel([22]), eliminates the necessity of separate consideration of functions of different numbers of variables. However it is an important assumption concerning $\mathfrak{S}$ that $\mathfrak{AB}$ is a term whenever $\mathfrak{A}$ and $\mathfrak{B}$ are, so that there must be terms which are intuitively meaningless.

3.123. Next it is fitting to point out, by a heuristic argument, that this involves no essential restriction on $\mathfrak{S}$. For suppose $\mathfrak{S}$ contains an enumerable (possibly infinite) set of operations $\mathfrak{O}_1, \mathfrak{O}_2, \cdots$ where $\mathfrak{O}_i$ has n_i arguments. With each $\mathfrak{O}_i$ let us associate a term $\mathfrak{O}_i^*$ (possibly a new primitive term); then if we replace $\mathfrak{O}_i(\mathfrak{A}_1, \cdots, \mathfrak{A}_{n_i})$ everywhere by $\mathfrak{O}_i^*\mathfrak{A}_1 \cdots \mathfrak{A}_{n_i}$ we have a new system $\mathfrak{S}^*$ in which the only operation is application. Of course the notion of term in $\mathfrak{S}^*$ is more general than in $\mathfrak{S}$; but a limitation to terms corresponding to terms in $\mathfrak{S}$ can be incorporated in the rules of procedure, and this limitation can presumably be expressed by elementary propositions. Suppose, now, that $\mathfrak{S}$ has the primitive terms $E_1, E_2, E_3, \cdots$. Let $\mathfrak{C}$ be a new primitive term and define

$$E_1 \equiv \mathfrak{CC}, \quad E_2 \equiv \mathfrak{C}E_1, \quad E_3 \equiv \mathfrak{C}E_2, \cdots;$$

then the resulting system has essentially the same content as the old, and has only one binary operation and one primitive term([23]).

3.13. There is a single unary predicate denoted by the Frege sign '$\vdash$.' Elementary propositions are thus of the form $\vdash\mathfrak{A}$, where $\mathfrak{A}$ is a term.

3.131. As to the intuitive meaning of this predicate, note that the sentence '$\vdash\mathfrak{A}$' expresses the fact that $\mathfrak{A}$ belongs to a certain category of terms, which category is defined recursively by the axioms and rules of procedure. We shall

([22]) C 304.1.

([23]) The possibility of reducing to one primitive term in the above manner is mentioned in Schönfinkel's paper cited above. Somewhat the same idea appears in various papers by Chwistek, who uses a single primitive term 'c' and a single binary operation denoted by a prefixed '*.'

call the members of this category *assertions* and, at times, read '$\vdash \mathfrak{A}$' as '$\mathfrak{A}$ is an assertion' or '$\mathfrak{A}$ is asserted.' Thus '$\vdash$' plays the same role as Hilbert's 'ist beweisbar' or Church's 'is provable'(24).

3.132. As in 3.123 we can show that this involves no essential restriction on $\mathfrak{S}$. For additional predicates can be expressed by means of $\vdash$ and terms (cf. the definition of equality in 3.2).

3.133. As to the meaning of '$\vdash_x$,' see 2.53.

3.14. There is a finite number of axioms. The terms asserted in these I shall denote by $A_0, A_1, \cdots, A_s$, so that the axioms are

$$\vdash A_i \qquad (i = 0, 1, \cdots, s).$$

Although in 2.13 'axiom' was defined as designating a proposition, it will be convenient to apply this name also to the A_i themselves; this is in accordance with current usage among logicians, who almost universally use 'axiom' to designate "formulas," i.e., terms.

3.15. There are a finite number of rules of procedure of the form

$$\vdash \mathfrak{X}_1 \,\&\, \cdots \,\&\, \vdash \mathfrak{X}_n \rightarrow_X \vdash \mathfrak{Y},$$

where $\mathfrak{X}_1, \cdots, \mathfrak{X}_n, \mathfrak{Y}$ are $(X_1, \cdots, X_m)$-terms such that every X_i which occurs in $\mathfrak{Y}$ also occurs in some $\mathfrak{X}_j$. (The restriction to rules of this character is not quite trivial, inasmuch as it implies that all conditions for significance are expressed by elementary propositions(25).)

3.16. *Remark.* It is sufficient to assume that $\mathfrak{S}$ contains a subsystem satisfying 3.14 and 3.15(26), i.e., that $\mathfrak{S}$ contains a finite number of elementary theorems of kind 3.14 and metatheorems of kind 3.15. For then, the proof of 15.4 will go through in the subsystem, and from this the inconsistency of the whole system follows as in 15.5.

---

(24) This usage of '$\vdash$' differs somewhat from that of Frege and the *Principia Mathematica*. There the symbol is placed before expressions which are already thought of as sentences, or at least as noun clauses, to indicate that the proposition concerned is true. Here '$\vdash$' is an intransitive verb; when placed before a noun the two together form a sentence; the proposition so denoted may or may not be taken as true according to the context.

(25) That the rules should be of the form given is a suggestion of Rosser (see Church C 359.7, footnote on p. 358); in his paper C 546.1 he showed how the Church system could be formulated as a system with rules of that character. The Rosser technique is also applicable to an $\mathfrak{L}$-system; this I plan to show in detail in a forthcoming paper. [This has since appeared; see J 6, pp. 41–53 (1941).]

(26) And, for that matter, 3.11. Naturally the other assumptions 3.2–3.6 must be satisfied by the subsystem. As to which of these will be satisfied automatically when they are satisfied for the system, that is a question which is not here gone into.

Kleene has pointed out to me that similar finiteness restrictions do not enter into the proof given by him and Rosser; because they first proved the inconsistency of a special system and then, in Theorem C, showed that more general systems include this as a subsystem. This part of the Kleene-Rosser argument can be superimposed on that given here to show that the finiteness restrictions of 3.14 and 3.15 are unnecessary. This argument can, presumably, be consider-

3.2. *Equality.* There exists a constant Q such that the relation of equality defined by

$$\mathfrak{X} = \mathfrak{Y} \equiv \vdash Q\mathfrak{X}\mathfrak{Y}$$

has the properties 3.21–3.27([27]). In consequence of these, 3.28 and 3.29 make convenient abbreviations.

3.21. $\mathfrak{A} = \mathfrak{A}$.

3.22. $\mathfrak{A} = \mathfrak{B} \rightarrow \mathfrak{B} = \mathfrak{A}$.

3.23. $\mathfrak{A} = \mathfrak{B} \,\&\, \mathfrak{B} = \mathfrak{C} \rightarrow \mathfrak{A} = \mathfrak{C}$.

3.24. $\mathfrak{A} = \mathfrak{B} \rightarrow \mathfrak{C}\mathfrak{A} = \mathfrak{C}\mathfrak{B}$.

3.25. $\mathfrak{A} = \mathfrak{B} \rightarrow \mathfrak{A}\mathfrak{C} = \mathfrak{B}\mathfrak{C}$.

3.26. $\mathfrak{A} = \mathfrak{B} \,\&\, \vdash \mathfrak{A} \rightarrow \vdash \mathfrak{B}$.

3.27. (General consequence of the foregoing.) If certain terms $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ occurring in the construction of a term $\mathfrak{X}$ are replaced by terms $\mathfrak{B}_1, \cdots, \mathfrak{B}_m$, thus forming a term $\mathfrak{Y}$, then

$$\mathfrak{A}_1 = \mathfrak{B}_1 \,\&\, \cdots \,\&\, \mathfrak{A}_m = \mathfrak{B}_m \rightarrow \mathfrak{X} = \mathfrak{Y}.$$

3.28. DEFINITION. $\mathfrak{X} = \mathfrak{Y} = \mathfrak{Z} . \equiv . \mathfrak{X} = \mathfrak{Y} \,\&\, \mathfrak{Y} = \mathfrak{Z}$. (It then follows, by 3.23, that $\mathfrak{X} = \mathfrak{Z}$.)

3.29. DEFINITION. $\vdash \mathfrak{X} = \vdash \mathfrak{Y} . \equiv . \vdash \mathfrak{X} \,\&\, \mathfrak{X} = \mathfrak{Y}$. (It then follows, by 3.26, that $\vdash \mathfrak{Y}$.)

3.3. *Combinatorial completeness.* Let $\mathfrak{S}'$ be an extension of $\mathfrak{S}$, and let $X_1, X_2, \cdots, X_n$ be indeterminates for $\mathfrak{S}'$. Let $\mathfrak{M}$ be any term of $\mathfrak{S}'(X_1, X_2, \cdots, X_n)$. Then there exists a term $\mathfrak{M}^*$ of $\mathfrak{S}'$ such that

$$\mathfrak{M}^* X_1 X_2 \cdots X_n = \mathfrak{M}.$$

This $\mathfrak{M}^*$ I shall denote, following Church, by the notation

$$\lambda^n X_1 X_2 \cdots X_n . \mathfrak{M}([28]).$$

The above I shall call combinatorial completeness in the strong sense. Church

---

ably simplified. All that it it necessary to do is show that any system $\mathfrak{S}$ satisfying all the hypotheses of this paper except the finiteness restrictions contains a subsystem $\mathfrak{S}'$ which satisfies all, and the particular $\mathfrak{S}'$ used by Kleene and Rosser is not the only one which will serve the purpose. It is, for instance, almost immediately evident that any such $\mathfrak{S}$ will contain a subsystem which is an $\mathfrak{L}$-system, and that an $\mathfrak{L}$-system satisfies all the hypotheses of this paper has been shown in my papers cited elsewhere. The point will not be discussed further here.

([27]) In terms of the notions introduced in 3.3, 3.4, and 3.6 such a Q can be defined by

$$Q \equiv \lambda^2 XY . F(\lambda Z . ZX)(\lambda Z . ZY) I$$

(cf. Rosser C 546.1); but this is a detail which will be passed over.

([28]) The period is to be interpreted as a bracket separating the complex operation '$\lambda^n X_1 \cdots X_n$' from its operand. The latter extends as far to the right as is consistent with indicated parentheses and other symbols of separation (cf. 2.45).

imposes the additional condition that $\mathfrak{M}$ must actually contain all of the $X_1, X_2, \cdots, X_n$; when this is imposed I shall speak of combinatorial completeness in the weak sense. In either case we have the following properties 3.31–3.35.

3.31. $(\lambda^n X_1 \cdots X_n . \mathfrak{M}) X_1 \cdots X_n = \mathfrak{M}$. (This follows immediately by the definition.)

3.32. If $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_n$ are any terms, then

$$(\lambda^n X_1 \cdots X_n . \mathfrak{M}) \mathfrak{A}_1 \cdots \mathfrak{A}_n = \begin{bmatrix} \mathfrak{A}_1, \cdots, \mathfrak{A}_n \\ X_1, \cdots, X_n \end{bmatrix} \mathfrak{M}.$$

(This follows by 2.61.)

3.33. If $Y_1, Y_2, \cdots, Y_n$ are also indeterminates for $\mathfrak{S}'$, then

$$\lambda^n Y_1 \cdots Y_n . \begin{bmatrix} Y_1, \cdots, Y_n \\ X_1, \cdots, X_n \end{bmatrix} \mathfrak{M} = \lambda^n X_1 \cdots X_n . \mathfrak{M}.$$

(This is a separate assumption, but its plausibility is suggested by the meaning attached to indeterminates in 2.5.)

3.34. If $\mathfrak{M} =_X \mathfrak{N}$, then

$$\lambda^n X_1 \cdots X_n . \mathfrak{M} = \lambda^n X_1 \cdots X_n . \mathfrak{N}(^{29}).$$

3.35. $\lambda^{n+1} X_0 X_1 \cdots X_n . \mathfrak{M} = \lambda X_0 (\lambda^n X_1 X_2 \cdots X_n . \mathfrak{M})$. (This shows that it is sufficient to assume combinatorial completeness for $n = 1$; but it is convenient to have the above properties to refer to for general n.)

3.4. *Some special constants.*

3.41. $I \equiv \lambda X . X$. Then I is the identity combinator. It has (by 3.32) the property that for any term $\mathfrak{A}$, $I\mathfrak{A} = \mathfrak{A}$.

3.42. $\mathfrak{S} \equiv \lambda^3 X Y Z . Y(X Y Z)$. Then $\mathfrak{S}$ is the successor function for natural numbers (cf. 3.44).

3.43. $K \equiv \lambda^2 X Y . X$. This K is called the cancellation combinator. Its definition (unlike those of I and $\mathfrak{S}$) requires combinatorial completeness in the strong sense. It has the property that for any terms $\mathfrak{A}$ and $\mathfrak{B}$

$$K\mathfrak{A}\mathfrak{B} = \mathfrak{A}.$$

$K\mathfrak{A}$ is thus, intuitively, the function having $\mathfrak{A}$ as its constant value. In the case of combinatorial completeness in the weak sense such constancy functions cannot be defined over the range of all terms (i.e., for $\mathfrak{B}$ unrestricted), but constancy functions over restricted ranges can.

3.44. The natural numbers are represented in the system by a sequence of terms $\mathfrak{Z}_0, \mathfrak{Z}_1, \mathfrak{Z}_2, \cdots$, such that

(29) This is assumed for the system $\mathfrak{S}$ and its term extensions, but not for systems obtained by adjoining axioms to such term extensions. With reference to such axiom extensions this property is invoked only when the premise is derivable without using the extraneous axioms.

$$\mathfrak{Z}_{n+1} \equiv \mathfrak{S}\,\mathfrak{Z}_n.$$

3.45. If the system is combinatorially complete in the strong sense, then we define

$$\mathfrak{Z}_0 \equiv KI,$$

and in that case we have, for any term $\mathfrak{G}$,

$$\mathfrak{Z}_0\mathfrak{G} = I, \qquad \mathfrak{Z}_1\mathfrak{G} = \lambda X.\mathfrak{G}X, \qquad \mathfrak{Z}_2\mathfrak{G} = \lambda X.\mathfrak{G}(\mathfrak{G}X),$$

and generally

$$\mathfrak{Z}_{n+1}\mathfrak{G} = \lambda X.\mathfrak{G}(\mathfrak{Z}_n\mathfrak{G}X);$$

so that, if $\mathfrak{G}$ is a function, $\mathfrak{Z}_n\mathfrak{G}$ is $\mathfrak{G}^n$, the nth iterate of $\mathfrak{G}$([30]).

3.5. *The Gödel representation.* With each constant $\mathfrak{A}$ there is associated a positive integer $\mathfrak{n}(\mathfrak{A})$ such that:

3.51. A number is assigned to only one term, i.e., if $\mathfrak{n}(\mathfrak{A}) = \mathfrak{n}(\mathfrak{B})$, then $\mathfrak{A}$ and $\mathfrak{B}$ are identical, and this means that they are not merely equal but are the same combination of the same primitive terms.

3.52. There exists a recursive function $\mu(x, y)$([31]) such that:

3.521. $\mathfrak{n}(\mathfrak{A}\mathfrak{B}) = \mu(\mathfrak{n}(\mathfrak{A}), \mathfrak{n}(\mathfrak{B}))$;

3.522. $x < \mu(x, y)$; $y < \mu(x, y)$;

3.523. If $\mu(x, y) = \mu(x', y')$, then $x = x'$ & $y = y'$([32]);

3.524. $\mu(x, y) \neq e_k$ where e_k is the number of a primitive term.

3.6. *Deducibility assumptions.* There exist two terms $\mathfrak{N}$ and F whose intuitive meanings are as follows: $\mathfrak{N}$ represents the property of being an integer, i.e., the proposition $\vdash \mathfrak{N}\mathfrak{A}$ is to mean the same as '$\mathfrak{A}$ is a natural number;' on the other hand F represents the notion of being a function, in the sense

([30]) This way of introducing the natural numbers is due to Church (C 359.6, p. 863).

([31]) A μ satisfying all conditions would be given by

$$\mu(x, y) = 2^x \cdot 3^y$$

provided the primitive terms are assigned numbers not of that form. Another such μ can be obtained from a recursive enumeration of number pairs (such as the $\sigma(x, y)$ of Hilbert-Bernays (C 507.1, p. 321) by skipping enough numbers to get in the e_k. Still another possibility is the following. Let application be denoted, for the time being, by a prefixed '*' (à la Chwistek), so that no parentheses are necessary. Assign to E_i the number $i+1$ and to * the number 0; then every term is assigned a unique number in the $(i+1)$-adic system; and conversely every such number with not too many 0's is assigned to a unique term (the number of prefixed 0's is uniquely determined). From this a function μ can be determined.

([32]) We must understand this to mean that for all x, y, x', y' either $\mu(x, y) \neq \mu(x', y')$, or $x = x'$ and $y = y'$. In the notation of §6.9 this is expressed by

$$\left| \delta_3(\mu(x, y), \mu(x', y')) \right| \cdot (\delta_3(x, x') + \delta_3(y, y')) = 0.$$

A similar remark applies to 3.524. Note that a part of 3.523 and 3.524 is a consequence of 3.51; thus, for 3.524, if $a = \mathfrak{n}(\mathfrak{A})$ and $b = \mathfrak{n}(\mathfrak{B})$, $\mu(a, b) \neq e_k$ follows by 3.51; but if a and b are not assigned to terms, 3.524 is an additional restriction on μ. These additional restrictions facilitate the proofs in §9, but are not essential.

that '$\vdash F\mathfrak{A}\mathfrak{B}\mathfrak{G}$' is to mean the same as '$\mathfrak{G}$ is a function on $\mathfrak{A}$ to $\mathfrak{B}$'([33]). These intuitive meanings are expressed by the following formal assumptions concerning them([34]):

3.61. $\vdash \mathcal{N}\mathfrak{Z}_0$, i.e., $\mathfrak{Z}_0$ is a natural number.

3.62. $\vdash F\mathcal{N}\mathcal{N}\mathfrak{S}$; i.e., $\mathfrak{S}$ is a function on $\mathcal{N}$ to $\mathcal{N}$, i.e., a numerical function.

3.63. $\vdash F\mathfrak{A}\mathfrak{B}\mathfrak{G} \,\&\vdash \mathfrak{A}X \rightarrow_X \vdash \mathfrak{B}(\mathfrak{G}X)$, i.e., if $\mathfrak{G}$ is a function on $\mathfrak{A}$ to $\mathfrak{B}$ and X is on the range $\mathfrak{A}$, then $\mathfrak{G}X$ is in $\mathfrak{B}$. This is the characteristic property of F according to its intuitive meaning. This rule will be sometimes referred to as "*Principle F.*"

*3.64([35]). (Mathematical induction.) If

$$\vdash \mathfrak{M}\mathfrak{Z}_0 \,\&\vdash F\mathfrak{M}\mathfrak{M}\mathfrak{S},$$

then

$$\vdash \mathcal{N}X \rightarrow_X \vdash \mathfrak{M}X.$$

3.65. (Property of deductive completeness.) If $\mathfrak{M}_1, \cdots, \mathfrak{M}_n, \mathfrak{A}, \mathfrak{B}, \mathfrak{G}$ are terms of an extension $\mathfrak{S}'$ of $\mathfrak{S}$ with respect to $Y_1, \cdots, Y_m$, and X is an indeterminate for $\mathfrak{S}'$, and if

$$\vdash \mathfrak{M}_1 \,\& \cdots \&\vdash \mathfrak{M}_n \,\&\vdash \mathfrak{A}X \rightarrow_{X,Y} \vdash \mathfrak{B}(\mathfrak{G}X);$$

then

$$\vdash \mathfrak{M}_1 \,\& \cdots \,\&\, \mathfrak{M}_n \rightarrow_Y \vdash F\mathfrak{A}\mathfrak{B}\mathfrak{G}([36]).$$

([33]) The postulation of F is equivalent to postulating a term representing formal implication. For if P^* represents formal implication, P^* and F are interdefinable, thus

$$F = \lambda^3 XYZ.P^*X(\lambda U.Y(ZU)),$$
$$P^* = \lambda^3 XY.FXYI.$$

(It is understood, of course, that in place of 3.63 the appropriate rule to go with P^* is

$$\vdash P^*\mathfrak{A}\mathfrak{B} \,\&\vdash \mathfrak{A}U \rightarrow_U \vdash \mathfrak{B}U.)$$

Many formal relations are simpler in terms of F.

([34]) Of the "assumptions" here stated it is evident that the really important ones are 3.63 and 3.65, expressing the intuitive meaning of F and deductive completeness respectively. The other properties 3.61, 3.62, and 3.64 will automatically hold for any system which is adequate for arithmetic at all. Incidentally it is not necessary to postulate them as independent assumptions. For if, using the Λ defined in §5.4, we adopted the definition

$$\mathcal{N} = \lambda X.F(\lambda Y.\Lambda(Y\mathfrak{Z}_0)(FYY\mathfrak{S}))(\lambda Y.YX)I,$$

they would become deducible as theorems. (The proof of this uses the full strength of 3.65, but is no more difficult than the theorems in §5.) When 3.64 is deduced in this way, it holds in the sense of deducibility (cf. the next footnote).

([35]) On the significance of the * in 3.64 see the footnote to 2.42. That there is no loss of generality in assuming 3.64 in this way is evident from the preceding footnote, since the properties of conjunction (5.42) are deducible without the use of 3.64.

([36]) The referee has called attention to the fact that, since we do not have conjunction in the system, we need this assumption for $n \geqq 2$ in order to derive 5.42 according to the method

4. **General discussion of the proof.** The derivation of the paradox depends on three principal theorems which may be roughly described as follows:

First, the system $\mathfrak{S}$ is strong enough to include all of recursive arithmetic, in the sense that every recursive arithmetic function has a representative in $\mathfrak{S}$, called its formalization, such that the properties of the recursive functions are reflected in corresponding theorems about their formalizations.

Second, we can construct a constant, here called T, which formalizes the Gödel representation; this means that if $a = \mathfrak{n}(\mathfrak{A})$ the proposition

$$\mathfrak{A} = T\mathfrak{Z}_a$$

is formally demonstrable as an elementary theorem in $\mathfrak{S}$.

Third, we can construct within the system a formal enumeration of the assertions [3.131]. Indeed by the aid of the above described T such an enumeration can be obtained from a formalization Θ of a primitive recursive function $\theta(x)$, which enumerates the Gödel numbers of those assertions. For this Θ we can further prove that

$$\vdash \mathfrak{N}X \rightarrow_X T(\Theta X).$$

These three theorems are established in Parts II, III, and IV respectively. On the basis of them it is shown in Part V that the Richard paradox can be set up in the system on much the same lines as in the introduction.

One or two general remarks may be interpolated here before we proceed.

The reader should note the difference between proving that

$$\vdash \mathfrak{M}\mathfrak{Z}_n \qquad (n = 0, 1, 2, \cdots)$$

and that

$$\vdash \mathfrak{N}X \rightarrow_X \vdash \mathfrak{M}X.$$

The first of these requires simply that we know that $\mathfrak{M}\mathfrak{Z}_n$ is an assertion for every non-negative integral value of n; the second requires the exhibition of a process for passing from the premise to the conclusion by the use of the rules of procedure for $\mathfrak{S}(X)$ only, X being treated as a constant. The second of these theorems is stronger than the first; if we have the second, the first follows by 5.27 and 2.61.

Again the reader should avoid being confused by the different senses of

---

given below. Of course, as soon as conjunction is introduced there is no increased generality in assuming a general n. If 5.42 were deduced by the method sketched in 5.6, then it would be sufficient to assume 3.65 for $n = 1$.

With this form of 3.65 it is evident that it holds not only for $\mathfrak{S}$ but also for any system obtained from $\mathfrak{S}$ by adjoining additional axioms (even if they contain additional terms). This is also true for 3.61–3.64; for 3.61 and 3.62 are elementary propositions, and 3.63 and 3.64 are deducibility properties in the sense of the footnote to 2.42. The same invariance with respect to extension also holds for various theorems in the sequel; this is not always explicitly indicated.

'induction,' 'recursion' and the like. These are sometimes used in an intuitive sense, sometimes as referring to a formal process related to $\mathfrak{S}$ (e.g., a use of 3.64) or the system $\mathfrak{r}$ of §6.

Part II. Formalization of recursive arithmetic

We now turn to the proof of the first of the three major theorems mentioned in §4. Some preliminary lemmas concerning special constants are proved in §5. Then in §6 ordinary recursive arithmetic is formulated precisely as a formal system $\mathfrak{r}$. This leads up to the proof of the final theorems in §7.

5. **Some special constants and their properties.** In this section we study some special functions useful in setting up recursive arithmetic within the system. Besides some properties of the elementary constants already postulated, we study the ordered pair D_2 and the recursion combinator R. It is necessary to include some generalizations of the deducibility postulates of 3.6.

5.1. *The F sequence.* Although functions of several variables can be regarded as special cases of functions of one variable, yet it saves cumbersome formulas to introduce an abbreviation for the term, definable in terms of F, which plays the same role for functions of n variables that F does for functions of one variable. This term is defined by induction thus:

$$F_0 \equiv I, \qquad F_1 \equiv F,$$
$$F_{n+1} \equiv \lambda^{n+2}XY_1 \cdots Y_nZ.FX(F_nY_1 \cdots Y_nZ).$$

It will be seen that for instance $\vdash F_2\mathfrak{ABCG}$ says that $\mathfrak{G}$ is on $\mathfrak{A}$ to $F\mathfrak{BC}$, and hence is a function on the composite range $\mathfrak{AB}$ to $\mathfrak{C}$.

5.11. Theorem. *For $m \geqq 1$, $n \geqq 0$,*

$$F_{m+n} \equiv \lambda^{m+n+1}X_1 \cdots X_mY_1 \cdots Y_nZ.F_mX_1 \cdots X_m(F_nY_1 \cdots Y_nZ).$$

Proof. For $m=1$, this is true by definition. Suppose the theorem is true for $m=k$. Then [3.31]

$$(1) \qquad F_{n+k}X_1 \cdots X_kY_1 \cdots Y_nZ = F_kX_1 \cdots X_k(F_nY_1 \cdots Y_nZ).$$

Therefore [by definition of F_{k+1}]

$$FU(F_{n+k}X_1 \cdots X_kY_1 \cdots Y_nZ) = FU(F_kX_1 \cdots X_k(F_nY_1 \cdots Y_nZ))$$
$$= F_{k+1}UX_1 \cdots X_k(F_nY_1 \cdots Y_nZ).$$

The theorem for $m=k+1$ now follows by 3.34.

*5.12. Theorem. (Generalization of Principle F.) *If, for $m \geqq 1$, $n \geqq 0$,*
(a) $\vdash F_{m+n}\mathfrak{K}_1\mathfrak{K}_2 \cdots \mathfrak{K}_m\mathfrak{L}_1\mathfrak{L}_2 \cdots \mathfrak{L}_n\mathfrak{MG}$, *and*
(b) $\vdash \mathfrak{K}_i\mathfrak{A}_i$ $(i=1, 2, \cdots, m)$;
then

$$\vdash F_n\mathfrak{L}_1\mathfrak{L}_2 \cdots \mathfrak{L}_n\mathfrak{M}(\mathfrak{G}\mathfrak{A}_1\mathfrak{A}_2 \cdots \mathfrak{A}_m).$$

Proof. Suppose that for a given $k<m$ we have established

$$\vdash F_{m+n-k}\mathfrak{K}_{k+1} \cdots \mathfrak{K}_m\mathfrak{L}_1 \cdots \mathfrak{L}_n\mathfrak{M}(\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_k). \tag{1}$$

Then [5.11] $\vdash F\mathfrak{K}_{k+1}(F_{m+n-k-1}\mathfrak{K}_{k+2} \cdots \mathfrak{K}_m\mathfrak{L}_1 \cdots \mathfrak{L}_n\mathfrak{M})(\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_k)$. Then by Principle F and the $k+1$st hypothesis (b) we have (1) for $k+1$ in the place of k. On the other hand (1) is true for $k=0$ by (a). Hence by induction (1) is true for all $k \leqq m$, hence for $k=m$, which gives the theorem.

5.13. THEOREM. (Generalization of 3.65.) *If*

$$\vdash \mathfrak{A}_1X_1 \,\&\, \cdots \,\&\, \vdash \mathfrak{A}_mX_m \rightarrow_X \vdash \mathfrak{B}(\mathfrak{G}X_1 \cdots X_m);$$

then

$$\vdash F_m\mathfrak{A}_1 \cdots \mathfrak{A}_m\mathfrak{B}\mathfrak{G}.$$

Proof. By induction from 3.65.

5.14. THEOREM. *If* (a) $\vdash F_n\mathfrak{A}_1 \cdots \mathfrak{A}_n\mathfrak{B}\mathfrak{G}$, (b) $\vdash F\mathfrak{A}_i'\mathfrak{A}_iI$ $(i=1, \cdots, n)$, (c) $\vdash F\mathfrak{B}\mathfrak{B}'I$; *then*

$$\vdash F_n\mathfrak{A}_1' \cdots \mathfrak{A}_n'\mathfrak{B}'\mathfrak{G}.$$

Proof. The proof follows by 5.13, since from the hypothesis

$$\vdash \mathfrak{A}_i'X_i \qquad (i = 1, 2, \cdots, n),$$

we conclude $\vdash \mathfrak{B}'(\mathfrak{G}X_1 \cdots X_n)$ by the hypotheses, 5.12, and 3.41.

5.2. *Properties of the elementary constants* I, K, $\mathfrak{S}$, $\mathfrak{Z}_n$.

5.21. THEOREM. $\vdash_U FUUI$.

Proof. By 3.41,

$$\vdash UX \rightarrow_{U,X} \vdash U(IX),$$

whence the theorem follows by 3.65.

5.22. THEOREM. $\vdash_U F_2U_1U_2U_1K$.

Proof. By 3.43,

$$\vdash U_1X_1 \,\&\, \vdash U_2X_2 \rightarrow_{U,X} \vdash U_1(KX_1X_2).$$

The theorem follows by 5.13.

5.23. THEOREM.

$$\vdash_U F(F(FU_1U_3)(FU_2U_1))(F(FU_1U_3)(FU_2U_3))\mathfrak{S}.$$

Proof. Suppose

$$\vdash U_2Z, \tag{1}$$

(2) $\vdash FU_1U_3Y,$

(3) $\vdash F_2(FU_1U_3)U_2U_1X.$

Then [5.12]

$$\vdash U_1(XYZ).$$

Therefore

$$\vdash U_3(Y(XYZ)) \quad \text{by [5.12]},$$

(4) $$= \vdash U_3(\mathfrak{S}XYZ) \quad [3.42; 3.2].$$

Thus (1), (2), (3) imply (4). Hence, by 5.13,

$$\vdash_V F_3(F_2(FU_1U_3)U_2U_1)(FU_1U_3)U_2U_3\mathfrak{S}.$$

By 5.11 this gives the theorem.

5.231. COROLLARY. $\vdash_V F(F(FUU)(FUU))(F(FUU)(FUU))\mathfrak{S}.$

5.24. THEOREM. $\vdash_V F_2U_1U_2U_2Z_0.$

Proof. Similar to that of 5.22; or from 5.22 (substituting FU_2U_2 for the U_1 and U_1 for the U_2 of 5.22) and 5.21 by Principle F.

5.25. THEOREM. $\vdash_V F(FUU)(FUU)\mathfrak{Z}_n$ $(n=0, 1, \cdots)$.

Proof. For $n=0$, this follows from 5.24 (take $U_1 \equiv FUU$, $U_2 \equiv U$). Assuming it for $\mathfrak{Z}_n$, it follows for $\mathfrak{Z}_{n+1}$ by 5.231, 3.63 (in view of 3.44). Hence the theorem follows by intuitive induction.

5.26. THEOREM. $\vdash_V F\mathfrak{N}(F(FUU)(FUU))I.$

Proof. Similar to that of 5.25; only we replace the intuitive induction by a formal induction using 3.64.

5.27. THEOREM. $\vdash \mathfrak{N}\mathfrak{Z}_n$ [3.61, 3.62, 3.63, 3.44].

5.3. *The ordered pair D_2.* The definition of the ordered pair here used, viz.,

$$D_2 \equiv \lambda^3XYZ.Z(KY)X,$$

is due to Bernays(37).

5.31. THEOREM. $D_2\mathfrak{A}\mathfrak{B}\mathfrak{Z}_0 = \mathfrak{A}.$

Proof. By 3.32

$$D_2\mathfrak{A}\mathfrak{B}\mathfrak{Z}_0 = \mathfrak{Z}_0(K\mathfrak{B})\mathfrak{A}$$
$$= \mathfrak{A} \quad [3.45, 3.41].$$

(37) Oral communication, May 2, 1936.

5.32. THEOREM. $D_2\mathfrak{A}\mathfrak{B}(\mathfrak{S}X)=\mathfrak{B}$.

Proof. By 3.32

$$\begin{aligned} D_2\mathfrak{A}\mathfrak{B}(\mathfrak{S}X) &= \mathfrak{S}X(K\mathfrak{B})\mathfrak{A} \\ &= K\mathfrak{B}(X(K\mathfrak{B})\mathfrak{A}) && [3.42; 3.32], \\ &= \mathfrak{B} && [3.43]. \end{aligned}$$

5.33. THEOREM. $\vdash_U F_3U_1U_2(F(FU_3U_2)(FU_1U_4))U_4D_2$.

Proof. If we assume $\vdash U_1X$, $\vdash U_2Y$, and $\vdash F(FU_3U_2)(FU_1U_4)Z$, then [5.22; 5.12] $\vdash FU_3U_2(KY)$. Therefore [5.12]

$$\vdash U_4(D_2XYZ).$$

The theorem follows by 5.13.

5.34. THEOREM. $\vdash_U F_3UUN(UD_2$.

Proof. In 5.33 put $U_1\equiv U_2\equiv U_3\equiv U_4\equiv U$. The theorem follows by 5.26, 5.21, and 5.14.

5.4. *Definition of conjunction.*

5.41. DEFINITION. $\Lambda\equiv\lambda^2XY.F(F(FII)(FII))I(D_2XY)$.

5.42. THEOREM. $\vdash\Lambda\mathfrak{A}\mathfrak{B}\rightleftarrows\vdash\mathfrak{A}\ \&\vdash\mathfrak{B}$.

Proof. Suppose $\vdash\Lambda\mathfrak{A}\mathfrak{B}$. Then, since by 5.25 $\vdash F(FII)(FII)\mathfrak{Z}_0$ and $\vdash F(FII)(FII)\mathfrak{Z}_1$, we have by Principle F

$$\vdash D_2\mathfrak{A}\mathfrak{B}\mathfrak{Z}_0\ \&\vdash D_2\mathfrak{A}\mathfrak{B}\mathfrak{Z}_1.$$

Therefore [5.31, 5.32]

(1) $$\vdash\mathfrak{A}\ \&\vdash\mathfrak{B}.$$

Conversely suppose (1) holds. Then by 3.41

(2) $$\vdash I\mathfrak{A}\ \&\vdash I\mathfrak{B}.$$

But, by 5.33,

(3) $$\vdash F_3II(F(FII)(FII))ID_2.$$

Therefore [(2); (3); 5.12]

$$\vdash F(F(FII)(FII))I(D_2\mathfrak{A}\mathfrak{B}) = \vdash\Lambda\mathfrak{A}\mathfrak{B},$$

q.e.d.

5.43. The more general conjunction Λ_n is definable recursively by

$$\Lambda_2\equiv\Lambda,$$

$$\Lambda_{n+1}\equiv\lambda^{n+1}XY_1\cdots Y_n.\Lambda X(\Lambda_nY_1\cdots Y_n).$$

It can then be shown, by intuitive induction, that

$$\vdash \Lambda_n \mathfrak{A}_1 \cdots \mathfrak{A}_n$$

is equivalent to

$$\vdash \mathfrak{A}_1 \,\&\, \cdots \,\&\, \vdash \mathfrak{A}_n.$$

5.44. THEOREM. (Generalization of mathematical induction.) *If*

(a) $\vdash \mathfrak{M}\mathsf{Z}_0$,

(b) $\vdash \mathcal{N}X \,\&\, \vdash \mathfrak{M}X \rightarrow_X \vdash \mathfrak{M}(\mathsf{S}X)$,

then

(c) $\vdash \mathcal{N}X \rightarrow_X \vdash \mathfrak{M}X$.

Proof. Let $\mathfrak{M}' \equiv \lambda X.\Lambda(\mathcal{N}X)(\mathfrak{M}X)$. Then [(a), 3.61]

$$\vdash \mathfrak{M}'\mathsf{Z}_0. \tag{1}$$

Assume

$$\vdash \mathfrak{M}'X. \tag{2}$$

Then [5.42] $\vdash \mathcal{N}X \,\&\, \vdash \mathfrak{M}X$. Therefore [(b)] $\vdash \mathfrak{M}(\mathsf{S}X)$. But [3.62, 3.63] $\vdash \mathcal{N}(\mathsf{S}X)$. Therefore [5.42] $\vdash \mathfrak{M}'(\mathsf{S}X)$. Since this follows from (2), we have by 3.65

$$\vdash F\mathfrak{M}'\mathfrak{M}'\mathsf{S}.$$

Therefore [3.64]

$$\vdash F\mathcal{N}\mathfrak{M}'I.$$

From this the theorem follows by 3.63, 3.41, and 5.42.

5.5. *The recursion combinator R.* We seek now to find a function $\mathfrak{G}$ which satisfies the following "definition" in terms of a given function $\mathfrak{H}$ and constant $\mathfrak{A}$

$$\mathfrak{G}\mathsf{Z}_0 = \mathfrak{A},$$

$$\mathcal{N}X \rightarrow_X \mathfrak{G}(\mathsf{S}X) = \mathfrak{H}X(\mathfrak{G}X).$$

This $\mathfrak{G}$ is defined by

$$\mathfrak{G} \equiv R\mathfrak{H}\mathfrak{A},$$

where R is the combinator defined in 5.52. 5.51 is preliminary; 5.52 contains the definition; while the theorems are stated in 5.53 and 5.54. The method of treatment is due to Bernays[38].

5.51. In order to find such a $\mathfrak{G}$ we consider first that if $\mathfrak{G}$ satisfied the condition

---

[38] See note to 5.3.

$$\mathfrak{G}(\mathfrak{S}X) = \mathfrak{H}(\mathfrak{G}X),$$

its definition would be very simple; viz.,

$$\mathfrak{G} \equiv \lambda X . X\mathfrak{H}\mathfrak{A}.$$

In the more difficult case suppose we had such a $\mathfrak{G}$. Then the function

$$\mathfrak{L} \equiv \lambda X . D_2 X(\mathfrak{G}X)$$

satisfies a recursion of the simpler type. For, assuming the above condition for $\mathfrak{G}$, we obtain

$$\begin{aligned}
\mathfrak{L}(\mathfrak{S}X) &= D_2(\mathfrak{S}X)(\mathfrak{G}(\mathfrak{S}X)) \\
&= D_2(\mathfrak{S}X)(\mathfrak{H}X(\mathfrak{G}X)) \\
&= D_2(\mathfrak{S}(\mathfrak{L}XZ_0))(\mathfrak{H}(\mathfrak{L}XZ_0)(\mathfrak{L}XZ_1)) \qquad [5.31, 5.32], \\
&= \mathfrak{K}(\mathfrak{L}X),
\end{aligned}$$

where $\mathfrak{K} \equiv \lambda X . D_2(\mathfrak{S}(XZ_0))(\mathfrak{H}(XZ_0)(XZ_1))$. This enables us to define $\mathfrak{L}$ and hence $\mathfrak{G}$.

These $\mathfrak{G}$, $\mathfrak{L}$ and $\mathfrak{K}$ are functions of $\mathfrak{H}$ and $\mathfrak{A}$. If we express them as such we have the R, $\mathfrak{R}_3$ and $\mathfrak{R}_2$ of the following definitions. It is convenient to define $\mathfrak{R}_2$ in two stages, of which $\mathfrak{R}_1$ is the first stage.

5.52. DEFINITIONS.

$$\mathfrak{R}_1 \equiv \lambda^3 UXY . D_2(\mathfrak{S}X)(UXY), \qquad \mathfrak{R}_2 \equiv \lambda^2 UX . \mathfrak{R}_1 U(XZ_0)(XZ_1),$$
$$\mathfrak{R}_3 \equiv \lambda^3 UVX . X(\mathfrak{R}_2 U)(D_2 Z_0 V), \qquad R \equiv \lambda^3 UVX . \mathfrak{R}_3 UVXZ_1.$$

5.53. THEOREM. *If we define* $\mathfrak{G} \equiv RUV$; *then*

(a) $\mathfrak{G}Z_0 =_{UV} V$,

(b) $\vdash \mathcal{N}X \rightarrow_{UVX} \mathfrak{G}(\mathfrak{S}X) = UX(\mathfrak{G}X)$.

Proof. Let us define

$$\mathfrak{K}_1 \equiv \mathfrak{R}_1 U, \tag{1}$$

$$\mathfrak{K}_2 \equiv \mathfrak{R}_2 U, \tag{2}$$

$$\mathfrak{L} \equiv \mathfrak{R}_3 UV, \tag{3}$$

$$\mathfrak{B} \equiv D_2 Z_0 V. \tag{4}$$

Then the above definitions give

$$\mathfrak{K}_1 = \lambda^2 XY . D_2(\mathfrak{S}X)(UXY), \tag{5}$$

$$\mathfrak{K}_2 = \lambda X . \mathfrak{K}_1(XZ_0)(XZ_1), \tag{6}$$

$$\mathfrak{L} = \lambda X . X\mathfrak{K}_2\mathfrak{B}, \tag{7}$$

$$\mathfrak{G} = \lambda X . \mathfrak{L}XZ_1. \tag{8}$$

Then [(7)]

$$\mathfrak{L}\mathfrak{Z}_0 = \mathfrak{Z}_0\mathfrak{K}_2\mathfrak{B}$$
$$= \mathfrak{B} \qquad [3.45].$$

Therefore

(9) $\mathfrak{L}\mathfrak{Z}_0\mathfrak{Z}_0 = \mathfrak{Z}_0$ [5.31],

(10) $\mathfrak{G}\mathfrak{Z}_0 = \mathfrak{L}\mathfrak{Z}_0\mathfrak{Z}_1 = V$ [5.32].

This proves the first part of the theorem. To prove the second part—

$$\mathfrak{L}(\mathfrak{S}X) = \mathfrak{S}X\mathfrak{K}_2\mathfrak{B} \qquad [(7)],$$
$$= \mathfrak{K}_2(X\mathfrak{K}_2\mathfrak{B}) \qquad [3.42],$$
$$= \mathfrak{K}_2(\mathfrak{L}X) \qquad [(7)],$$
$$= \mathfrak{K}_1(\mathfrak{L}X\mathfrak{Z}_0)(\mathfrak{G}X) \qquad [(6), (8)],$$

(11) $$= D_2(\mathfrak{S}(\mathfrak{L}X\mathfrak{Z}_0))(U(\mathfrak{L}X\mathfrak{Z}_0)(\mathfrak{G}X)) \qquad [(5)].$$

Therefore [5.31]

$$\mathfrak{L}(\mathfrak{S}X)\mathfrak{Z}_0 = \mathfrak{S}(\mathfrak{L}X\mathfrak{Z}_0).$$

This with (9) gives a proof by mathematical induction [5.44 with $\mathfrak{M} \equiv \lambda X . Q(\mathfrak{L}X\mathfrak{Z}_0)X$] that

$$\vdash \mathcal{N}X \to_X \mathfrak{L}X\mathfrak{Z}_0 = X.$$

If we put this in (11), then

$$\vdash \mathcal{N}X \to_X \mathfrak{G}(\mathfrak{S}X) = \mathfrak{L}(\mathfrak{S}X)\mathfrak{Z}_1$$
$$= UX(\mathfrak{G}X) \qquad [5.32],$$

q.e.d.

5.54. THEOREM. $\vdash F_2(F_2\mathcal{N}\mathfrak{Q}\mathfrak{Q})\mathfrak{Q}(F\mathcal{N}\mathfrak{Q})R$.

Proof. Let $\mathfrak{G}$ be as in 5.53, and suppose that

(1) $\vdash F_2\mathcal{N}\mathfrak{Q}\mathfrak{Q}U$,

(2) $\vdash \mathfrak{Q}V$.

Then [5.53 (a)] $\vdash \mathfrak{Q}(\mathfrak{G}\mathfrak{Z}_0)$.

Assume

(3) $\vdash \mathcal{N}X$ & $\vdash \mathfrak{Q}(\mathfrak{G}X)$.

Then [(1), 5.12] $\vdash \mathfrak{Q}(UX(\mathfrak{G}X))$; that is [5.53 (b)],

(4) $\vdash \mathfrak{Q}(\mathfrak{G}(\mathfrak{S}X))$.

Therefore [(3), (4), 5.44, 3.65]

$$\vdash F\mathfrak{N}\mathfrak{Q}\mathfrak{G}.$$

This means that, assuming (1) and (2), it follows that

$$\vdash F\mathfrak{N}\mathfrak{Q}(RUV).$$

The theorem follows by 5.13.

5.6. *Discussion.* Throughout the foregoing repeated use has been made of 3.65. Some interest attaches to the question of whether these theorems can be derived without using 3.65 in its full power, inasmuch as systems with weakened forms of 3.65 may conceivably be consistent. This section will be closed with some informal discussion of results bearing on this question.

Suppose we assume the following postulates concerning F:

$$\vdash F_2(FYZ)(FXY)(FXZ)B,$$
$$\vdash F(F_2YXZ)(F_2XYZ)C,$$
$$\vdash F(F_2XXY)(FXY)W,$$
$$\vdash FX(FYX)K,$$

where B, C, W are the "primitive combinators" which may be defined, in our present notation, as follows:

$$B \equiv \lambda^3XYZ.X(YZ),$$
$$C \equiv \lambda^3XYZ.XZY,$$
$$W \equiv \lambda^2XY.XYY.$$

These postulates may, of course, be expressed as axioms; they are easy consequences of 3.65. Moreover they are highly plausible assumptions concerning F (cf. the intuitive discussions in my paper C 396.7). With these postulates in the place of 3.65 all the above theorems except 5.13 can be derived, provided that mathematical induction in the form 5.44, and also 3.61, 3.62, and 5.26 be assumed([39]). This, however, I have accomplished not for the above definition of R, but for R defined in a way more akin to that used by Kleene. The sequence of definitions for this purpose is the following:

([39]) These derivations are contained in papers now in preparation which employed the methods of my paper J 1, 65. The essential point is the following lemma: If $\mathfrak{X}$ is an $(X_1, \cdots, X_n)$-term and $\mathfrak{A}_1, \cdots, \mathfrak{A}_n, \mathfrak{B}$ are constants such that from the assumptions

$$\vdash \mathfrak{A}_iX_i \qquad (i = 1, 2, \cdots, n),$$

it follows, by following through the construction of $\mathfrak{X}$ and applying Principle F only, that

$$\vdash \mathfrak{B}\mathfrak{X};$$

then

$$\vdash F_n\mathfrak{A}_1 \cdots \mathfrak{A}_n\,\mathfrak{B}(\lambda^nX_1 \cdots X_n.\mathfrak{X}).$$

This lemma can then replace 5.13. Presumably then the inductive assumptions can be derived by adjoining appropriate postulates for Λ, and defining $\mathfrak{N}$ as in the second footnote to 3.6.

$$\mathfrak{D}_1 \equiv \lambda X.D_2X(\mathfrak{S}X), \qquad \mathfrak{D}_2 \equiv \lambda X.\mathfrak{D}_1(XZ_1),$$
$$\mathfrak{D}_3 \equiv \lambda X.X\mathfrak{D}_2(D_2Z_0Z_0), \qquad \nabla \equiv \lambda X.\mathfrak{D}_3XZ_0,$$
$$\mathfrak{R}_1 \equiv \lambda^3UYX.UX(Y(\nabla X)), \qquad \mathfrak{R}_2 \equiv \lambda^3UVX.X(\mathfrak{R}_1U)(KV),$$
$$R \equiv \lambda^3UVX.\mathfrak{R}_2UVX(\nabla X).$$

The ∇ so defined is the predecessor function; it has the properties

$$\nabla Z_0 = Z_0, \qquad \vdash \mathcal{N}X \to_X \nabla(\mathfrak{S}X) = X,$$

while R has the same properties as in 5.53 and 5.54. The new definitions are easier to handle because the ordered pairs involved are homogeneous, whereas in 5.52 they are heterogeneous. On the present basis, however, the Bernays definition is easier.

Note, too, that the definition of conjunction in §5.4 is such as not to require any use of mathematical induction (even in the above weakened system). If the definition

$$\lambda^2XY.F\mathcal{N}I(D_2XY)$$

were used, then we should require 5.34, and hence 5.26, to establish 5.42.

Many of the theorems in the later sections can also be derived in the so-modified system. How far we can go with it is, as yet, undetermined.

6. **Recursive arithmetic as a formal system.** We are now in possession of enough formal machinery to show that recursive arithmetic can be incorporated in the system $\mathfrak{S}$. In order to make this result explicit, it is desirable to formulate recursive arithmetic with all possible precision. In the literature recursive arithmetic is usually treated in a semi-intuitive manner. However, greater exactness and definiteness are attained by treating it as itself a formal system. The present section will be devoted to formulating recursive arithmetic in that manner. After some preliminary explanations in 6.1, the formal statement of the primitive frame of the system is given in 6.2–6.7; in 6.8 there are some general theorems about the system $\mathfrak{r}$, and in 6.9 there are stated, for future reference, the definitions of some special recursive functions([40]).

6.1. *Preliminary explanations.* The system of recursive arithmetic will be called the system $\mathfrak{r}$. Its terms will be called *expressions* to distinguish them from the terms of the system $\mathfrak{S}$. These expressions include not only numbers and numerical variables, but also functions of every degree of multiplicity. Accordingly we consider an infinite series of categories of expressions $\mathfrak{e}_0, \mathfrak{e}_1, \mathfrak{e}_2, \cdots$. The expressions of $\mathfrak{e}_0$ are the numerical variables and expressions capable of being substituted for them—we shall call these *numerical expressions*—those of $\mathfrak{e}_n$ for $n>0$ are *functions* of n arguments. It is also

([40]) See also my paper *A formalization of recursive arithmetic*, American Journal of Mathematics, vol. 63 (1941), pp. 263–282. (This paper was written after the presentation.)

desirable to make explicit the variables on which a numerical expression depends or may depend; thus if $\mathfrak{x}$ is any set of numerical variables $\mathfrak{e}_0(\mathfrak{x})$ is the set of all numerical expressions whose variables, if any, are in $\mathfrak{x}$. In this connection the null set of variables is denoted by $\mathfrak{o}$ and the set of all variables by $\mathfrak{v}$, so that $\mathfrak{e}_0$ is the same as $\mathfrak{e}_0(\mathfrak{v})$. On the use of letters for expressions see 2.72. '$\in$' and '$\subset$' are used with their ordinary set-theoretic meanings. In the statement of the primitive frame parts of the text enclosed in parentheses refer to the intuitive interpretation or other extraneous consideration; these are inserted to increase intelligibility, but are not a part of the formal specifications.

6.2. *Primitive expressions.* There is an infinite set of these, constituted as follows:

6.21. (A constant numeral) 0.

6.22. An infinite sequence $\mathfrak{v}$ of *numerical variables*: $v_1, v_2, v_3, \cdots$. The letters 'x,' 'y,' 'z,' etc., stand for unspecified members of this sequence.

6.23. (A function) σ (this is the successor function, converting any numerical expression $\mathfrak{a}$ into $\mathfrak{a}+1$).

6.24. (A function) κ_0 (this is the function having 0 as its constant value; cf. 6.62 below).

6.25. A set (of functions) κ_{nm} where $n=1, 2, 3, \cdots$ and $m=1, 2, 3, \cdots, n$. (κ_{nm} picks out the mth argument from a sequence of n; cf. 6.63 below.)

6.3. *Operations.* These also form an infinite set, but may be grouped into three kinds, as in 6.31–6.33. Operations are denoted by Gothic letters, which are prefixed to their arguments; since the multiplicity of each is definite, parentheses are unnecessary (as in the logical notation of the Poles).

6.31. For each $n>0$, there is an $(n+1)$-ary operation $\mathfrak{a}_n$. (This represents the application of an n-ary function to n arguments. In accordance with custom) let us define

$$\mathfrak{g}(\mathfrak{a}_1, \mathfrak{a}_2, \cdots, \mathfrak{a}_n) \equiv \mathfrak{a}_n\mathfrak{g}\mathfrak{a}_1\mathfrak{a}_2 \cdots \mathfrak{a}_n,$$

when $\mathfrak{g}\in\mathfrak{e}_n$ and $\mathfrak{a}_1, \mathfrak{a}_2, \cdots, \mathfrak{a}_n\in\mathfrak{e}_0$.

6.32. For each $n>0$, and $m>0$ there is an $(m+1)$-ary operation $\mathfrak{s}_{mn}$ (representing the substitution of m n-ary functions in an m-ary one, cf. 6.64 below).

6.33. For each $n\geqq 0$ a binary $\mathfrak{r}_n$ (representing the recursive definition of an $(n+1)$-ary function by an n-ary one and an $(n+2)$-ary one, cf. 6.65 below).

6.4. *Rules of formation.* We turn now to the rules for the formation of derived expressions, and for the classification of expressions into the categories mentioned in 6.1.

6.41. $0\in\mathfrak{e}_0(\mathfrak{x})$ for any $\mathfrak{x}$.

6.42. $\mathfrak{x}\subset\mathfrak{e}_0(\mathfrak{x})$.

6.43. $\sigma\in\mathfrak{e}_1$.

6.44. $\kappa_0\in\mathfrak{e}_1$.

6.45. $\kappa_{nm}\in\mathfrak{e}_n$.

6.46. If $\mathfrak{g} \in \mathfrak{e}_n$ and $\mathfrak{a}_i \in \mathfrak{e}_0(\mathfrak{x})$ for $i = 1, 2, \cdots, n$, then $\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n) \in \mathfrak{e}_0(\mathfrak{x})$.

6.47. If $\mathfrak{h} \in \mathfrak{e}_m$ and $\mathfrak{k}_i \in \mathfrak{e}_n$ for $i = 1, 2, \cdots, m$, then $\mathfrak{S}_{mn}\mathfrak{h}\mathfrak{k}_1 \cdots \mathfrak{k}_m \in \mathfrak{e}_n$.

6.48. If $\mathfrak{h} \in \mathfrak{e}_n$ and $\mathfrak{k} \in \mathfrak{e}_{n+2}$, then $\mathfrak{R}_n\mathfrak{h}\mathfrak{k} \in \mathfrak{e}_{n+1}$[41].

6.5. *Elementary propositions.* The only primitive predicate of the system is the binary relation of equality between numerical expressions. The elementary propositions are those of the form

$$\mathfrak{a} = \mathfrak{b},$$

where $\mathfrak{a} \in \mathfrak{e}_0(\mathfrak{v})$, $\mathfrak{b} \in \mathfrak{e}_0(\mathfrak{v})$.

6.6. *Axioms.* In stating the axioms we have two alternatives: either we may state specific axioms with a substitution rule, or we may state axiom-schemes in the sense of von Neumann. The latter course is followed here. It is, then, to be understood that $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ are any expression of $\mathfrak{e}_0$ and $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{k}$ are functions.

6.61. $\mathfrak{a} = \mathfrak{a}$[42].

6.62. $\mathfrak{K}_0(\mathfrak{a}) = 0$.

6.63. $\mathfrak{K}_{nm}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n) = \mathfrak{a}_m$.

6.64. If $\mathfrak{h} \in \mathfrak{e}_m$ and $\mathfrak{k}_1, \cdots, \mathfrak{k}_m \in \mathfrak{e}_n$ and $\mathfrak{g} \equiv \mathfrak{S}_{mn}\mathfrak{h}\mathfrak{k}_1 \cdots \mathfrak{k}_m$, then

$$\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n) = \mathfrak{h}(\mathfrak{k}_1(\mathfrak{a}_1, \cdots, \mathfrak{a}_n), \cdots, \mathfrak{k}_m(\mathfrak{a}_1, \cdots, \mathfrak{a}_n)).$$

6.65. If $\mathfrak{h} \in \mathfrak{e}_n$[43], $\mathfrak{k} \in \mathfrak{e}_{n+2}$, $\mathfrak{g} \equiv \mathfrak{R}_n\mathfrak{h}\mathfrak{k}$, then

$$\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, 0) = \mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n),$$

$$\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \sigma(\mathfrak{b})) = \mathfrak{k}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \mathfrak{b}, \mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \mathfrak{b})).$$

6.7. *Rules of procedure.* As before $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ are arbitrary expressions of $\mathfrak{e}_0$.

6.71. If $\mathfrak{a} = \mathfrak{b}$, then $\mathfrak{b} = \mathfrak{a}$.

6.72. If $\mathfrak{a} = \mathfrak{b}$, and $\mathfrak{b} = \mathfrak{c}$, then $\mathfrak{a} = \mathfrak{c}$.

6.73. If $\mathfrak{g} \in \mathfrak{e}_n$, and if

$$\mathfrak{a}_i = \mathfrak{b}_i \qquad (i = 1, 2, \cdots, n);$$

then $\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n) = \mathfrak{g}(\mathfrak{b}_1, \cdots, \mathfrak{b}_n)$.

6.74. (Mathematical induction[44].) If $\mathfrak{g}$, $\mathfrak{h} \in \mathfrak{e}_{n+1}$, $\mathfrak{a}_1, \cdots, \mathfrak{a}_n \in \mathfrak{e}_0(\mathfrak{x})$ are such that

(a) $\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, 0) = \mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, 0)$,

(b) on the hypothesis that

(1) $$\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, y) = \mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, y),$$

---

[41] If $n = 0$, then $\mathfrak{e}_n$ is to be understood as meaning $\mathfrak{e}_0(\mathfrak{o})$.

[42] This is, of course, redundant.

[43] If $n = 0$, $\mathfrak{h}$ must be in $\mathfrak{e}_0(\mathfrak{o})$.

[44] It will be noticed that this rule has a metatheoretic premise, and therefore does not have the simple character prescribed for the rules of $\mathfrak{S}$ in 3.14.

where y is a variable not in $\mathfrak{x}$, it follows by 6.71–6.73 (and known theorems) that

$$\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \sigma(y)) = \mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \sigma(y)); \tag{2}$$

then $\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \mathfrak{b}) = \mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \mathfrak{b})$.

6.8. *Some general theorems.* The following theorems are obvious enough from the intuitive point of view; but they may be proved to follow formally from the above definitions. The proofs are omitted here; it is planned to publish them in a separate paper(48).

6.81. THEOREM. *If* $\mathfrak{a} \in \mathfrak{e}_0(x_1, \cdots, x_m)$ *and* $\mathfrak{a}'$ *is obtained from* $\mathfrak{a}$ *by substituting for* $x_1, \cdots, x_m$ *expressions* $\mathfrak{b}_1, \cdots, \mathfrak{b}_m \in \mathfrak{e}_0(y_1, \cdots, y_n)$, *then* $\mathfrak{a}' \in \mathfrak{e}_0(y_1, \cdots, y_n)$.

6.82. THEOREM. *If* $\mathfrak{a}, \mathfrak{b} \in \mathfrak{e}_0(x_1, \cdots, x_n)$ *such that*

$$\mathfrak{a} = \mathfrak{b};$$

and if $\mathfrak{a}', \mathfrak{b}'$ *are obtained by substituting* $\mathfrak{c}_1, \cdots, \mathfrak{c}_n \in \mathfrak{e}_0$ *for* $x_1, \cdots, x_n$ *respectively in* $\mathfrak{a}$ *and* $\mathfrak{b}$; *then*

$$\mathfrak{a}' = \mathfrak{b}'.$$

6.83. THEOREM. *If* $\mathfrak{a} \in \mathfrak{e}_0(x_1, \cdots, x_n)$, *then there exists a* $\mathfrak{g} \in \mathfrak{e}_n$ *such that*

$$\mathfrak{a} = \mathfrak{g}(x_1, \cdots, x_n).$$

This $\mathfrak{g}$ *we denote by* $\mathfrak{f}_{x_1 \cdots x_n}(\mathfrak{a})$; *it is defined by intuitive recursion as follows*:

$$\mathfrak{f}_{x_1 \cdots x_n}(x_i) \equiv \mathfrak{k}_{ni},$$

$$\mathfrak{f}_{x_1 \cdots x_n}(0) \equiv \mathfrak{s}_{1n}\mathfrak{k}_0\mathfrak{k}_{n1},$$

$$\mathfrak{f}_{x_1 \cdots x_n}(\mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_m)) \equiv \mathfrak{s}_{mn}\mathfrak{h}\mathfrak{g}_1 \cdots \mathfrak{g}_m,$$

where $\mathfrak{g}_i \equiv \mathfrak{f}_{x_1 \cdots x_n}(\mathfrak{a}_i)$.

6.84. THEOREM. *If* $\mathfrak{a} \in \mathfrak{e}_0(x_1, \cdots, x_n)$, $\mathfrak{b} \in \mathfrak{e}_0(x_1, \cdots, x_n, y, z)$; *then there exists a* $\mathfrak{g} \in \mathfrak{e}_{n+1}$ *such that*

$$\mathfrak{g}(x_1, \cdots, x_n, 0) = \mathfrak{a}, \qquad \mathfrak{g}(x_1, \cdots, x_n, \sigma(y)) = \mathfrak{b}',$$

where $\mathfrak{b}'$ *is the result of substituting* $\mathfrak{g}(x_1, \cdots, x_n, y)$ *for* z *in* $\mathfrak{b}$.

6.85. THEOREM. (Generalization of mathematical induction.) *If* $\mathfrak{a}$, $\mathfrak{b}$ *are in* $\mathfrak{e}_0(x_1, \cdots, x_n, y)$ *such that*

(a) *we have*

$$\left[\begin{matrix}0\\y\end{matrix}\right]\mathfrak{a} = \left[\begin{matrix}0\\y\end{matrix}\right]\mathfrak{b},$$

(48) See the paper cited in the footnote to the beginning of §6.

(b) *from the assumption*

$$\mathfrak{a} = \mathfrak{b}$$

it follows by rules 6.71–6.73 *only that*

$$\begin{bmatrix} \sigma(y) \\ y \end{bmatrix}\mathfrak{a} = \begin{bmatrix} \sigma(y) \\ y \end{bmatrix}\mathfrak{b};$$

then

$$\mathfrak{a} = \mathfrak{b}.$$

6.86. The formulation above given is sufficient for recursive arithmetic as defined by Hilbert and Bernays([46]), in the sense that, if

$$\mathfrak{a} = \mathfrak{b}$$

is deducible in the formalism of Hilbert-Bernays, then it is deducible also in the system r.

6.9. *Special functions.* The object of this section is to give explicitly, primarily for the purpose of fixing the notation, the definitions of some special recursive functions which we need in the sequel. For some of these functions we use an operational, rather than a Greek letter notation. The functions are defined within the system r by virtue of 6.83 or 6.84. Some of the properties of these functions are also listed. On the proofs of the statements made see 6.8 (especially 6.86).

For explanation of references to this number in the later portions of the paper see 7.6.

6.91. The sum $x+y$ is defined by

$$x + 0 = x, \qquad x + \sigma(y) = \sigma(x + y),$$

and obeys the commutative and associative laws.

6.92. The product $x \cdot y$ is defined by

$$x \cdot 0 = 0, \qquad x \cdot \sigma(y) = x \cdot y + x.$$

This product is commutative and associative and distributive with respect to addition.

6.93. Functions $\delta_1(x)$ and $x \dot{-} y$, related to subtraction, are defined as follows:

$$\delta_1(0) = 0, \qquad \delta_1(\sigma(x)) = x;$$
$$x \dot{-} 0 = x, \qquad x \dot{-} \sigma(y) = \delta_1(x \dot{-} y).$$

(Intuitively $x \dot{-} y$ is $x - y$ if $x \geqq y$, otherwise it is 0.) Some of the properties of these functions are as follows:

([46]) C 507.1, p. 307.

6.931. $\delta_1(a) = a \dot{-} 1$.
6.932. $a \dot{-} a = 0 \dot{-} a = 0$.
6.933. $(a+c) \dot{-} (b+c) = a \dot{-} b$.
6.934. $(a+b) \dot{-} b = a$.
6.935. $a \dot{-} (b+c) = (a \dot{-} b) \dot{-} c = (a \dot{-} c) \dot{-} b$.
6.936. $a \cdot (b \dot{-} c) = (a \cdot b) \dot{-} (a \cdot c)$.
6.937. $(a \dot{-} b) \cdot (b \dot{-} a) = 0$.
6.938. $a + (b \dot{-} a) = b + (a \dot{-} b)$.

6.94. In terms of subtraction we define the negation function $|x|$ thus:

$$|x| = 1 \dot{-} x.$$

Then we have the following properties:
6.941. $|0| = 1$; $|\sigma(a)| = 0$.
6.942. $||0|| = 0$; $||\sigma(a)|| = 1$.
6.943. If $|a| \cdot b = 0$, and $a = 0$, then $b = 0$.
6.944. $a = ||a|| + \delta_1(a)$.

6.95. The equality function δ_2 is defined by

$$\delta_2(x, y) = (x \dot{-} y) + (y \dot{-} x).$$

6.96. General sums and products are defined thus:
(a) $\sum_{y=0}^{x} g(y)$ is the function $h(x)$ such that

$$h(0) = g(0), \qquad h(\sigma(x)) = h(x) + g(\sigma(x)).$$

(b) $\prod_{y=0}^{x} g(y)$ is the function $f(x)$ such that

$$f(0) = g(0), \qquad f(\sigma(x)) = f(x) \cdot g(\sigma(x)).$$

The first of these functions is 0 if and only if $g(y) = 0$ for every $y \leqq x$; the second is 0 if and only if $g(y) = 0$ for some $y \leqq x$. It is also convenient to define

$$(\exists y \leqq x) g(y) = \left|\left| \prod_{y=0}^{x} g(y) \right|\right|.$$

6.97. Given $g(x)$, $(\epsilon y \leqq x) g(y)$ is defined as that function $h(x)$ such that

$$h(0) = 0,$$

$$h(\sigma(x)) = h(x) \cdot \left| \prod_{y=0}^{x} g(y) \right| + \sigma(x) \cdot |\, g(\sigma(x))| \cdot (\exists y \leqq x) g(y).$$

Intuitively $(\epsilon y \leqq x) g(y)$ is the least $z \leqq x$ such that $g(z) = 0$, if there is any such z; otherwise it is 0. Thus we have the following properties:
6.971. $(\epsilon y \leqq x) g(y) \dot{-} x = 0$.
6.972. $|(\exists y \leqq x) g(y)| \cdot g((\epsilon y \leqq x) g(y)) = 0$.

6.98. Given a, $b_0, \cdots, b_p$, $c \in e_0(x_1, \cdots, x_n)$, a function g such that

$$\mathfrak{g}(x_1, \cdots, x_n) = \begin{cases} b_k & \text{if } \mathfrak{a} = k \leqq p, \\ c & \text{if } |\mathfrak{a} \dot{-} p| = 0, \end{cases}$$

is given by

$$\mathfrak{g}(x_1, \cdots, x_n) \equiv \sum_{k=0}^{p} |\delta_2(\mathfrak{a}, k)| \cdot b_k + \|\mathfrak{a} \dot{-} p\| \cdot c.$$

6.99. The function

$$\left[\begin{matrix} x \\ y \end{matrix}\right],$$

representing as usual the integral part of x/y, may be defined thus:

$$\left[\begin{matrix} x \\ y \end{matrix}\right] \equiv (\epsilon z \leqq x)(\sigma(x) \dot{-} y \cdot \sigma(z)).$$

7. **The representation of recursive functions.** In this section we study a method of associating with every expression $\mathfrak{m}$, numerical or functional, of the system $\mathfrak{r}$, a term called its formalization and denoted by '$\mathfrak{F}(\mathfrak{m})$.' This association is such that if $\mathfrak{m}$ is a function, or is an $\mathfrak{e}_0(\mathfrak{o})$, then $\mathfrak{F}(\mathfrak{m})$ is a constant in $\mathfrak{S}$; while if $\mathfrak{m}$ contains variables, $\mathfrak{F}(\mathfrak{m})$ contains corresponding indeterminates. Furthermore, with each elementary proposition of $\mathfrak{r}$ we associate a proposition concerning $\mathfrak{S}$ which we call the formalization of the former proposition; if this elementary proposition is $\mathfrak{a} = \mathfrak{b}$, and $X_1, X_2, \cdots, X_n$ are the indeterminates in $\mathfrak{F}(\mathfrak{a})$ and $\mathfrak{F}(\mathfrak{b})$, this formalization is

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_n \rightarrow_X \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b}).$$

The object is to show that this formalization is true whenever the original proposition of $\mathfrak{r}$ is true.

The definition of $\mathfrak{F}$ is given in 7.1, and some preliminary theorems are found in 7.2; in 7.3 it is shown that the formalizations of all the $\mathfrak{r}$-axioms are true, while in 7.4 we establish the validity of the formalized $\mathfrak{r}$-rules as principles of inference between such formalized propositions; this leads to the general result and some related general theorems in 7.5. In 7.6 are listed some special cases which are useful in what follows.

7.1. *Definition of* $\mathfrak{F}$. $\mathfrak{F}(\mathfrak{m})$ is defined by recursion on the construction of $\mathfrak{m}$ as follows:

7.11. $\mathfrak{F}(0) \equiv Z_0$.

7.12. $\mathfrak{F}(v_i) \equiv V_i$—that is, to each variable v_i of $\mathfrak{r}$ we associate a particular indeterminate V_i. When we use 'x,' 'x_k,' 'y_j,' etc., to stand for unspecified variables of $\mathfrak{r}$, then it is to be understood that 'X,' 'X_k,' 'Y_j,' etc., stand for the corresponding indeterminates of $\mathfrak{S}$.

7.13. $\mathfrak{F}(\sigma) \equiv S$.

7.14. $\mathfrak{F}(\kappa_0) \equiv KZ_0$.

7.15. $\mathfrak{F}(\kappa_{nm}) \equiv \lambda^n X_1 \cdots X_n . X_m$.

7.16. If $\mathfrak{F}(\mathfrak{g}) \equiv \mathfrak{G}$, and $\mathfrak{F}(\mathfrak{a}_i) \equiv \mathfrak{A}_i$, then

$$\mathfrak{F}(\mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n)) \equiv \mathfrak{G}\mathfrak{A}_1\mathfrak{A}_2 \cdots \mathfrak{A}_n.$$

7.17. If $\mathfrak{g} \equiv \mathfrak{s}_{mn}\mathfrak{h}\mathfrak{k}_1 \cdots \mathfrak{k}_m$, and if $\mathfrak{F}(\mathfrak{h}) \equiv \mathfrak{H}$, $\mathfrak{F}(\mathfrak{k}_i) \equiv \mathfrak{K}_i$, then

$$\mathfrak{F}(\mathfrak{g}) \equiv \lambda^n X_1 \cdots X_n . \mathfrak{H}(\mathfrak{K}_1 X_1 \cdots X_n) \cdots (\mathfrak{K}_m X_1 \cdots X_n).$$

7.18. If $\mathfrak{h} \equiv \mathfrak{r}_n\mathfrak{h}\mathfrak{k}$, and $\mathfrak{F}(\mathfrak{h}) \equiv \mathfrak{H}$, $\mathfrak{F}(\mathfrak{k}) \equiv \mathfrak{K}$, then

$$\mathfrak{F}(\mathfrak{g}) \equiv \lambda^{n+1} X_1 \cdots X_n Y . R\mathfrak{K}^*\mathfrak{H}Y X_1 \cdots X_n,$$

where

$$\mathfrak{K}^* \equiv \lambda^{n+2} Y Z X_1 \cdots X_n . \mathfrak{K} X_1 \cdots X_n Y(Z X_1 \cdots X_n).$$

7.2. *Functional character of the formalization of a numerical function.* It is convenient to introduce $\mathcal{N}_k$ as an abbreviation for $F_k\mathcal{N} \cdots \mathcal{N}$ (where the '$\mathcal{N}$' is repeated $k+1$ times). The precise definition is a recursive one, thus (cf. 5.1):

$$\mathcal{N}_0 \equiv \mathcal{N}, \qquad \mathcal{N}_{k+1} \equiv F\mathcal{N}\mathcal{N}_k.$$

7.21. THEOREM. *If* $\mathfrak{g} \in \mathfrak{e}_n$, *and* $\mathfrak{G} \equiv \mathfrak{F}(\mathfrak{g})$, *then* $\vdash \mathcal{N}_n\mathfrak{G}$.

Proof. By induction on the construction of $\mathfrak{g}$ in 7.211–7.215.

7.211. The theorem is true for $\mathfrak{g} \equiv \sigma$ by 7.13, and 3.62.

7.212. For $\mathfrak{g} \equiv \kappa_0$ the theorem follows by 7.14, 5.22 (with $U_1 \equiv U_2 \equiv \mathcal{N}$), 3.61, and 5.12.

7.213. The theorem is true for $\mathfrak{g} \equiv \kappa_{nm}$, for

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_n \to_X \vdash \mathcal{N}X_m \qquad (m \leqq n),$$
$$\to_X \vdash \mathcal{N}(\mathfrak{F}(\kappa_{nm}))X_1 \cdots X_n \qquad [7.15, 3.31, 3.26],$$

whence the theorem follows by 5.13.

7.214. Suppose $\mathfrak{g} \equiv \mathfrak{s}_{mn}\mathfrak{h}\mathfrak{k}_1, \cdots, \mathfrak{k}_m$, and the theorem is true for $\mathfrak{h}, \mathfrak{k}_1, \cdots, \mathfrak{k}_m$. Let $\mathfrak{G} \equiv \mathfrak{F}(\mathfrak{g})$, $\mathfrak{H} \equiv F(\mathfrak{h})$, $\mathfrak{K}_j \equiv \mathfrak{F}(\mathfrak{k}_j)$ $(j = 1, 2, \cdots, m)$. Then by the hypothesis of the induction and 5.12

(1) $\vdash \mathcal{N}Y_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}Y_m \to_Y \vdash \mathcal{N}(\mathfrak{H}Y_1 \cdots Y_m)$,

$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_n \to_X \vdash \mathcal{N}(\mathfrak{K}_j X_1 \cdots X_n) \qquad (j = 1, \cdots, m),$

$\to \vdash \mathcal{N}(\mathfrak{H}(\mathfrak{K}_1 X_1 \cdots X_n) \cdots (\mathfrak{K}_m X_1 \cdots X_n))$

$[(1), 2.61],$

$\to \vdash \mathcal{N}(\mathfrak{G}X_1 \cdots X_n) \qquad [7.17, 3.31, 3.26].$

The theorem then follows by 5.13.

7.215. Suppose $\mathfrak{g} \equiv \mathfrak{r}_n\mathfrak{h}\mathfrak{k}$, and that the hypothesis of the induction is true for $\mathfrak{h}$ and $\mathfrak{k}$. Let $\mathfrak{H} \equiv \mathfrak{F}(\mathfrak{h})$, $\mathfrak{K} \equiv \mathfrak{F}(\mathfrak{k})$, and $\mathfrak{K}^*$ be as in 7.18. Then by the hypothesis of the induction,

(1) $$\vdash \mathcal{N}_n\mathfrak{H},$$

and [5.11]

(2) $$\vdash F_2\mathcal{N}\mathcal{N}\mathcal{N}_n\mathfrak{K}.$$

From (2) and 5.12, 5.13 (much as in 7.214) we can then conclude

(3) $$\vdash F_2\mathcal{N}\mathcal{N}_n\mathcal{N}_n\mathfrak{K}^*.$$

From (1), (3), 5.54, and 5.12 we have

(4) $$\vdash F\mathcal{N}\mathcal{N}_n(R\mathfrak{K}^*\mathfrak{H}).$$

But, by 7.18,

$$\mathfrak{G} \equiv \lambda^{n+1}X_1 \cdots X_n Y.R\mathfrak{K}^*\mathfrak{H}YX_1 \cdots X_n.$$

Hence from (4), using first 5.12 then 5.13, we have the theorem for $\mathfrak{G}$.

7.22. THEOREM. *If* $\mathfrak{a} \in \mathfrak{e}_0(x_1, \cdots, x_m)$, *then*

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow \vdash \mathcal{N}(\mathfrak{F}(\mathfrak{a})).$$

Proof. We can prove the theorem by induction on the construction of $\mathfrak{a}$, bearing in mind that if $\mathfrak{a}$ is primitive then it must be 0 or some v_i, and if it is not primitive it must be constructed from primitives by 6.46 only. If $\mathfrak{a} = 0$, the result follows by 7.11 and 3.61. If $\mathfrak{a} = v_i$, the theorem to be proved becomes $\vdash \mathcal{N}V_i \rightarrow \vdash \mathcal{N}V_i$ (by 7.12) which is a tautology. Suppose $\mathfrak{a} \equiv \mathfrak{g}(\mathfrak{a}_1, \cdots, \mathfrak{a}_n)$ and that the theorem is true for $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$. Let $\mathfrak{A}_i \equiv \mathfrak{F}(\mathfrak{a}_i)$, $\mathfrak{A} \equiv \mathfrak{F}(\mathfrak{a})$, $\mathfrak{G} \equiv \mathfrak{F}(\mathfrak{g})$. Then, by the hypothesis of the induction,

$$\begin{aligned} \vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m &\rightarrow \vdash \mathcal{N}\mathfrak{A}_j && (j = 1, 2, \cdots, n), \\ &\rightarrow \vdash \mathcal{N}(\mathfrak{G}\mathfrak{A}_1, \cdots, \mathfrak{A}_n) && [7.21, 5.12], \\ &\rightarrow \vdash \mathcal{N}\mathfrak{A} && [7.16], \end{aligned}$$

q.e.d.

7.3. *Verification of the axioms.* We are now in a position to show that if $\mathfrak{a} = \mathfrak{b}$ is an axiom, and $\mathfrak{a}, \mathfrak{b} \in \mathfrak{e}_0(x_1, \cdots, x_m)$, then $\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_x \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b})$.

7.31. In the case of the axiom-schemes 6.61–6.64 we have

$$\vdash \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b})$$

without any hypothesis [by 3.21; 3.43; 3.32; respectively]; hence, what is to be proved follows a fortiori.

7.32. Let us now turn to the axiom-schemes of 6.65. Suppose that $\mathfrak{a}_1, \cdots, \mathfrak{a}_n, \mathfrak{b} \in \mathfrak{e}_0(x_1, \cdots, x_m)$, and let $\mathfrak{A}_j \equiv \mathfrak{F}(\mathfrak{a}_j)$, $\mathfrak{B} \equiv \mathfrak{F}(\mathfrak{b})$, $\mathfrak{G} \equiv \mathfrak{F}(\mathfrak{g})$, $\mathfrak{H} \equiv \mathfrak{F}(\mathfrak{h})$, $\mathfrak{K} \equiv \mathfrak{F}(\mathfrak{k})$. Then by 7.13, 7.16, what we have to show is that if

(1) $$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m,$$

then

$$\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n \mathfrak{Z}_0 = \mathfrak{H}\mathfrak{A}_1 \cdots \mathfrak{A}_n, \tag{2}$$

$$\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n(\mathfrak{S}\mathfrak{B}) \equiv \mathfrak{K}\mathfrak{A}_1 \cdots \mathfrak{A}_n\mathfrak{B}(\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n\mathfrak{B}). \tag{3}$$

Let $\mathfrak{K}^*$ be defined as in 7.18, and let $\mathfrak{G}^* \equiv R\mathfrak{K}^*\mathfrak{H}$. Then [7.18, 3.32]

$$\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n Y =_Y \mathfrak{G}^* Y \mathfrak{A}_1 \cdots \mathfrak{A}_n. \tag{4}$$

By 5.53, $\mathfrak{G}^*\mathfrak{Z}_0 = \mathfrak{H}$. From this and (4) we have (2).

Again, by 5.53,

$$\vdash \mathcal{N}Y \rightarrow_Y \mathfrak{G}^*(\mathfrak{S}Y) = \mathfrak{K}^* Y(\mathfrak{G}^* Y). \tag{5}$$

By 7.2, on the hypothesis (1) we have

$$\vdash \mathcal{N}\mathfrak{B}.$$

Therefore [(5), 2.61]

$$\mathfrak{G}^*(\mathfrak{S}\mathfrak{B}) = \mathfrak{K}^*\mathfrak{B}(\mathfrak{G}^*\mathfrak{B}); \tag{6}$$

hence

$$\begin{aligned}
\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n(\mathfrak{S}\mathfrak{B}) &= \mathfrak{G}^*(\mathfrak{S}\mathfrak{B})\mathfrak{A}_1 \cdots \mathfrak{A}_n && [(4)],\\
&= \mathfrak{K}^*\mathfrak{B}(\mathfrak{G}^*\mathfrak{B})\mathfrak{A}_1 \cdots \mathfrak{A}_n && [(6)],\\
&= \mathfrak{K}\mathfrak{A}_1 \cdots \mathfrak{A}_n\mathfrak{B}(\mathfrak{G}^*\mathfrak{B}\mathfrak{A}_1 \cdots \mathfrak{A}_n) && [7.18, 3.32],\\
&= \mathfrak{K}\mathfrak{A}_1 \cdots \mathfrak{A}_n\mathfrak{B}(\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n\mathfrak{B}) && [(4)].
\end{aligned}$$

Then (3) follows by 3.23.

7.4. *Verification of the rules.* Next we have to show that the rules of procedure of $\mathfrak{r}$ are valid principles of inference between the formalizations of the elementary propositions.

7.41. As a preliminary to this let us make a remark concerning these formalizations. The formalization of $\mathfrak{a} = \mathfrak{b}$ is defined as

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b}), \tag{1}$$

where $X_1, \cdots, X_m$ are the indeterminates, if any, corresponding to variables appearing in $\mathfrak{a}$ or $\mathfrak{b}$ (or both), and hence are all the indeterminates in $\mathfrak{F}(\mathfrak{a})$ or $\mathfrak{F}(\mathfrak{b})$ (or both). But this formalization is equivalent to

$$\vdash \mathcal{N}Y_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}Y_n \rightarrow_Y \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b}), \tag{2}$$

provided $Y_1, \cdots, Y_n$ are any set of indeterminates which include the $X_1, \cdots, X_m$. For if (1) is true, then (2) is true a fortiori. On the other hand, if (2) is true, then for any Y_i not occurring in either $\mathfrak{F}(\mathfrak{a})$ or $\mathfrak{F}(\mathfrak{b})$ substitute $\mathfrak{Z}_0$ for Y_i; the corresponding premise is redundant by 3.61, and the conclusion

is unaltered; the premises can then be rearranged (the relation designated by '&' is intuitive conjunction) to give (1).

7.42. This established, we may show that the formalized 6.71–6.73 follow from the rules of 3.2. Let us take 6.72 as a typical case; the others are proved similarly. In accordance with 7.41 what we have to show is that, if

(1) $$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b}),$$

(2) $$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \mathfrak{F}(\mathfrak{b}) = \mathfrak{F}(\mathfrak{c}),$$

then

(3) $$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{c}),$$

where $X_1, \cdots, X_m$ is a set of indeterminates including all which occur in $\mathfrak{F}(\mathfrak{a})$, $\mathfrak{F}(\mathfrak{b})$, or $\mathfrak{F}(\mathfrak{c})$. Very well, let (1) and (2) be assumed. Then on the assumption that the hypothesis of (3) holds, we have $\mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b})$ and $\mathfrak{F}(\mathfrak{b}) = \mathfrak{F}(\mathfrak{c})$, hence $\mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{c})$ by 3.23.

7.43. Let us turn now to the rule of mathematical induction (6.74). Let $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ be as in the hypotheses of 6.74, and let $\mathfrak{G} \equiv \mathfrak{F}(\mathfrak{g})$, $\mathfrak{H} \equiv \mathfrak{F}(\mathfrak{h})$, $\mathfrak{A}_i \equiv \mathfrak{F}(\mathfrak{a}_i)$ $(i = 1, 2, \cdots, n)$, $\mathfrak{B} \equiv \mathfrak{F}(\mathfrak{b})$. Let $X_1, \cdots, X_m$ be a sufficiently inclusive set of indeterminates, and let

$$\mathfrak{K} \equiv Q(\mathfrak{G}\mathfrak{A}_1 \cdots \mathfrak{A}_n Y)(\mathfrak{H}\mathfrak{A}_1 \cdots \mathfrak{A}_n Y),$$
$$\mathfrak{L} \equiv \lambda^{n+1} Y X_1 \cdots X_m . \mathfrak{K}.$$

Then the formalization of hypothesis (a), which we assume to hold, is

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \vdash \mathfrak{L}Z_0 X_1 \cdots X_m,$$

which is equivalent [3.41] to

(1) $$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \vdash I(\mathfrak{L}Z_0 X_1 \cdots X_m).$$

If, now, we define

$$\mathfrak{M} \equiv \lambda Y . F_m \mathcal{N} \cdots \mathcal{N} I(\mathfrak{L}Y),$$

then from (1), 5.13 we have

(2) $$\vdash \mathfrak{M}Z_0.$$

Suppose, now, that

(3) $$\vdash \mathcal{N}Y \,\&\, \vdash \mathfrak{M}Y.$$

Then [5.12, definition of $\mathfrak{M}$, 3.41]

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \vdash \mathfrak{L}Y X_1 \cdots X_m.$$

Therefore, a fortiori,

(4) $$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \,\&\, \vdash \mathcal{N}Y \rightarrow_{X,Y} \vdash \mathfrak{L}Y X_1 \cdots X_m.$$

This is the formalization of 6.74 (1). Now the passage 6.74 (1) to 6.74 (2) was by rules 6.71–6.73 only; hence, by 7.42 from (4) we can pass([47]) to

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \,\&\, \vdash \mathcal{N}Y \rightarrow_{X,Y} \vdash \mathfrak{L}(\mathfrak{S}\,Y)X_1 \cdots X_m,$$

and thence [(3)] to

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow_X \vdash \mathfrak{L}(\mathfrak{S}\,Y)X_1 \cdots X_m.$$

Therefore [5.13]

$$\vdash \mathfrak{M}(\mathfrak{S}\,Y). \tag{5}$$

We have deduced (5) on the hypothesis (3); hence by (2) and 5.44

$$\vdash \mathcal{N}Y \rightarrow_Y \vdash \mathfrak{M}Y.$$

Therefore [5.12]

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \,\&\, \vdash \mathcal{N}Y \rightarrow \vdash \mathfrak{R}.$$

The formalization of the conclusion of 6.74 now follows by substitution of $\mathfrak{B}$ for Y, which is permissible by 2.61 and 7.22.

7.5. *General theorems.* As mentioned above, the following may be established.

7.51. THEOREM. *If*

$$\mathfrak{a} = \mathfrak{b} \tag{1}$$

is valid in $\mathfrak{r}$, *and if* $X_1, \cdots, X_n$ *are all the indeterminates occurring in* $\mathfrak{F}(\mathfrak{a})$ *or* $\mathfrak{F}(\mathfrak{b})$, *then*

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_n \rightarrow_X \mathfrak{F}(\mathfrak{a}) = \mathfrak{F}(\mathfrak{b}). \tag{2}$$

Proof. By induction on the proof of (1) in 7.3 and 7.4.

7.52. THEOREM. *If* $\mathfrak{a} \in \mathfrak{e}_0(x_1, \cdots, x_n)$, *and if we define*

$$\mathfrak{F}_{x_1 x_2 \cdots x_n}(\mathfrak{a}) \equiv \mathfrak{F}(\mathfrak{f}_{x_1 \cdots x_n}(\mathfrak{a})),$$

where $\mathfrak{f}_{x_1 \cdots x_n}(\mathfrak{a})$ *is defined as in* 6.83, *then*

([47]) In case known theorems are used in 6.74 we must imagine that we are dealing with a particular proof of a theorem in $\mathfrak{r}$, and showing that this can be converted into a proof in $\mathfrak{S}$ of the corresponding formalization. If we suppose the theorems in the original proof are numbered consecutively from 1 to n, then we have to show by induction on k that the formalization of the first k steps are theorems in $\mathfrak{S}$. Supposing this done for a given k, we are interested here in the case where the $k+1$st step is by 6.74. Then the known theorems used will be those of index $\leq k$. For these the formalizations in $\mathfrak{S}$ are true by the hypothesis of the induction. The argument in 7.42 shows that the formalization of 6.74 (2) is deducible from (4) and these formalizations, and hence, since the latter are true, from (4) only. This deducibility depends, of course, on properties of $\rightarrow$ not explicitly formulated in this paper; but we may suppose, by 7.41, that all the formalizations in this subsidiary proof have the same premises, and then the deducibility relations hold between the conclusions—as in 7.42—by virtue of the rules of procedure of $\mathfrak{S}$ alone.

$$\mathfrak{F}_{x_1 \cdots x_n}(\mathfrak{a}) = \lambda^n X_1 \cdots X_n . \mathfrak{F}(\mathfrak{a}).$$

Proof. We proceed by intuitive induction in the construction of $\mathfrak{a}$. We adopt the abbreviation $\mathfrak{g} \equiv \mathfrak{f}_{x_1 \cdots x_n}(\mathfrak{a})$.

If $\mathfrak{a} \equiv x_i$, then $\mathfrak{F}(\mathfrak{a}) \equiv X_i$, $\mathfrak{g} \equiv \kappa_{ni}$ and the theorem holds by 7.15. If $\mathfrak{a} \equiv 0$, then $\mathfrak{g} \equiv \mathfrak{s}_{1n}\kappa_0\kappa_{n1}$, and if we set

$$\mathfrak{K}_{nm} \equiv \mathfrak{F}(\kappa_{nm}),$$

we have [7.17, 7.14]

$$\begin{aligned} \mathfrak{F}(\mathfrak{g}) &\equiv \lambda^n X_1 \cdots X_n . K \mathfrak{Z}_0 (\mathfrak{K}_{n1} X_1 \cdots X_n), \\ &= \lambda^n X_1 \cdots X_n . \mathfrak{Z}_0 && [3.34, 3.43], \\ &= \lambda^n X_1 \cdots X_n . \mathfrak{F}(\mathfrak{a}) && [7.11]. \end{aligned}$$

The theorem is therefore true if $\mathfrak{a}$ is primitive.

Suppose $\mathfrak{a} \equiv \mathfrak{h}(\mathfrak{a}_1, \cdots, \mathfrak{a}_m)$, and the theorem is true for $\mathfrak{a}_1, \cdots, \mathfrak{a}_m$. Then, if we let $\mathfrak{g}_1, \cdots, \mathfrak{g}_m$ be as in 6.83, and set

$$\mathfrak{G}_i \equiv \mathfrak{F}(\mathfrak{g}_i), \qquad \mathfrak{A}_i \equiv \mathfrak{F}(\mathfrak{a}_i), \qquad \mathfrak{H} \equiv \mathfrak{F}(\mathfrak{h}),$$

we have

$$\mathfrak{G}_i \equiv \lambda^n X_1 \cdots X_n . \mathfrak{A}_i.$$

Therefore [3.31]

$$\mathfrak{G}_i X_1 \cdots X_n = \mathfrak{A}_i. \tag{1}$$

Again [6.83], $\mathfrak{g} \equiv \mathfrak{s}_{mn}\mathfrak{h}\mathfrak{g}_1 \cdots \mathfrak{g}_m$. Therefore

$$\begin{aligned} \mathfrak{F}(\mathfrak{g}) &\equiv \lambda^n X_1 \cdots X_n . \mathfrak{H}(\mathfrak{G}_1 X_1 \cdots X_n) \cdots (\mathfrak{G}_m X_1 \cdots X_n) && [7.17], \\ &= \lambda^n X_1 \cdots X_n . \mathfrak{H}\mathfrak{A}_1 \cdots \mathfrak{A}_m && [(1), 3.34], \\ &= \lambda^n X_1 \cdots X_n . \mathfrak{F}(\mathfrak{a}) && [7.16]. \end{aligned}$$

This completes the proof by intuitive induction.

7.53. THEOREM. *If a is a numeral*[48], *then* $\mathfrak{F}(a) = \mathfrak{Z}_a$.

Proof. Clear by 7.11, 7.13, 3.44.

7.54. THEOREM. *If $\mathfrak{g} \in \mathfrak{e}_n$, and $a_1, a_2, \cdots, a_n, b$ are numerals such that*

$$b = \mathfrak{g}(a_1, \cdots, a_n),$$

then

$$\mathfrak{Z}_b = \mathfrak{G}\mathfrak{Z}_{a_1}\mathfrak{Z}_{a_2} \cdots \mathfrak{Z}_{a_n},$$

where $\mathfrak{G} \equiv \mathfrak{F}(\mathfrak{g})$.

(48) A numeral is an expression of the form $\sigma(\sigma(\cdots \sigma(0) \cdots))$.

Proof. Clear by 7.51, 7.53, 7.16.

7.6. *Special functions.* The formalizations of the special functions of 6.9 will be symbolized as explained below. To substantiate statements of properties of these functions which follow by 6.9 and 7.5 I shall often refer simply to 6.9. In interpreting the operational definitions it is to be understood that the definitions hold when $\mathfrak{A}$, $\mathfrak{B}$, etc., are respectively the formalizations of $\mathfrak{a}$, $\mathfrak{b}$, etc.

$$\begin{aligned}
\mathfrak{A} + \mathfrak{B} &= \mathfrak{F}(\mathfrak{a} + \mathfrak{b}),\\
\mathfrak{A} \cdot \mathfrak{B} &= \mathfrak{F}(\mathfrak{a} \cdot \mathfrak{b}),\\
\Delta_1 &= \mathfrak{F}(\delta_1),\\
\mathfrak{A} \dot{-} \mathfrak{B} &= \mathfrak{F}(\mathfrak{a} \dot{-} \mathfrak{b}),\\
|\mathfrak{A}| &= \mathfrak{F}(|\mathfrak{a}|),\\
\Delta_2 &= \mathfrak{F}(\delta_2),\\
\sum_{k=0}^{p} \mathfrak{A}_k &= \mathfrak{F}\left(\sum_{k=0}^{p} \mathfrak{a}_k\right) (49).
\end{aligned}$$

Certain properties of these functions, useful as lemmas, are given in 7.61–7.63. In these $\mathfrak{A}$, $\mathfrak{B}$ stand for terms subject to the hypothesis $\vdash \mathcal{N}\mathfrak{A}$ & $\vdash \mathcal{N}\mathfrak{B}$.

7.61. THEOREM. $\mathfrak{A}+\mathfrak{B}=Z_0 \rightarrow \mathfrak{A}=Z_0 \;\&\; \mathfrak{B}=Z_0$.

Proof. By 6.932

$$\begin{aligned}
&\mathfrak{A} + \mathfrak{B} = Z_0 \rightarrow (\mathfrak{A} + \mathfrak{B}) \dot{-} \mathfrak{B} = Z_0;\\
(1)\qquad &\qquad\qquad \rightarrow \mathfrak{A} = Z_0 && [6.934].\\
&\mathfrak{A} + \mathfrak{B} = Z_0 \rightarrow \mathfrak{B} + \mathfrak{A} = Z_0 && [6.91],\\
&\qquad\qquad\quad \rightarrow \mathfrak{B} = Z_0 && [(1)].
\end{aligned}$$

7.62. THEOREM. $\Delta_2\mathfrak{A}\mathfrak{B}=Z_0 \rightarrow \mathfrak{A}=\mathfrak{B}$.

Proof. We have

$$\begin{aligned}
\Delta_2\mathfrak{A}\mathfrak{B} = Z_0 &\rightarrow (\mathfrak{A} \dot{-} \mathfrak{B}) + (\mathfrak{B} \dot{-} \mathfrak{A}) = Z_0 && [\text{definition}]\\
&\rightarrow \mathfrak{A} \dot{-} \mathfrak{B} = Z_0 \;\&\; \mathfrak{B} \dot{-} \mathfrak{A} = Z_0 && [7.61],\\
&\rightarrow \mathfrak{A} = \mathfrak{A} + (\mathfrak{B} \dot{-} \mathfrak{A}) && [6.91],\\
&\qquad = \mathfrak{B} + (\mathfrak{A} \dot{-} \mathfrak{B}) && [6.938],\\
&\qquad = \mathfrak{B} && [6.91],
\end{aligned}$$

q.e.d.

*7.63. THEOREM. *If* $|\mathfrak{A}| \cdot \mathfrak{B} = Z_0$, *and* $\mathfrak{A} = Z_0$, *then* $\mathfrak{B} = Z_0$ [6.941, 6.92, 6.91].

(49) Note that p and k are here taken in an intuitive sense, and are not formalized.

PART III. FORMALIZATION OF THE GÖDEL REPRESENTATION

The subject of this part is the second of the major theorems considered in §4, the construction of T. §8 is preliminary; the construction proper is given in §9.

8. Numerical functions associated with the Gödel representation.

8.1. *The functions* $\mu_1(x)$, $\mu_2(x)$, $\mu_3(x)$. We seek now primitive recursive functions μ_1 and μ_2 such that if $x = \mathfrak{n}(\mathfrak{X}_1\mathfrak{X}_2)$ then $\mu_1(x) = \mathfrak{n}(\mathfrak{X}_1)$, $\mu_2(x) = \mathfrak{n}(\mathfrak{X}_2)$. This we may do as follows:

$$\mu_1(x) \equiv (\epsilon y \leqq x)(\exists z \leqq x)\delta_2(x, \mu(y, z)), \tag{1}$$

$$\mu_2(x) \equiv (\epsilon z \leqq x)\delta_2(x, \mu(\mu_1(x), z)). \tag{2}$$

Then we have

$$\mu_1(\mu(x, y)) = x, \tag{3}$$

$$\mu_2(\mu(x, y)) = y. \tag{4}$$

If we define further

$$\mu_3(x) \equiv (\exists y \leqq x)(\exists z \leqq x)(\delta_2(x, \mu(y, z))), \tag{5}$$

then by the processes of recursive arithmetic we have

$$\mu_3(\mu(x, y)) = 0. \tag{6}$$

Also, $\mu_3(x) = 0 \rightarrow x = \mu(\mu_1(x), \mu_2(x))$; in fact, we have

$$\mu_3(x) = \|\delta_2(x, \mu(\mu_1(x), \mu_2(x)))\|. \tag{7}$$

Likewise we have

$$\mu_3(x) \cdot \mu_1(x) = \mu_3(x) \cdot \mu_2(x) = 0. \tag{8}$$

8.2. *The functions* M, M_1, M_2, M_3. These are defined as the formalizations of μ, μ_1, μ_2, μ_3, respectively. Then by 8.1 and the theorems of 7.5 we have the following.

8.21. $\vdash \mathcal{N}X \;\&\; \vdash \mathcal{N}Y \rightarrow_{X,Y} M_1(MXY) = X \;\&\; M_2(MXY) = Y$.

8.22. $\vdash \mathcal{N}X \;\&\; M_3X = \mathfrak{Z}_0 \rightarrow_X M(M_1X)(M_2X) = X$.

8.23. $\vdash \mathcal{N}X \;\&\; \vdash \mathcal{N}Y \rightarrow_{X,Y} M_3(MXY) = \mathfrak{Z}_0$.

8.3. *The function* $\mathfrak{T}$[50]. It is convenient to denote by '$\mathfrak{T}$' the correspondence inverse to $\mathfrak{n}$, so that

$$a = \mathfrak{n}(\mathfrak{A}) \rightleftarrows \mathfrak{T}(a) = \mathfrak{A}. \tag{1}$$

This defines $\mathfrak{T}$ uniquely for a number of the form $\mathfrak{n}(\mathfrak{A})$. For other values of x we may define

---

[50] This discussion of $\mathfrak{T}$ is for heuristic purposes only. The conclusions of the paper do not depend on the properties of $\mathfrak{T}$ here stated.

$$\mathfrak{T}(x) \equiv I.$$

Such a $\mathfrak{T}$ will have, then, the following property. Suppose $a = \mathfrak{n}(\mathfrak{A})$, $b = \mathfrak{n}(\mathfrak{B})$, $c = \mathfrak{n}(\mathfrak{C})$, where $\mathfrak{C} \equiv \mathfrak{A}\mathfrak{B}$. Then

$$c = \mu(a, b),$$
$$\mathfrak{C} = \mathfrak{T}(c) = \mathfrak{T}(\mu(a, b)),$$
$$\mathfrak{A}\mathfrak{B} = \mathfrak{T}(a)(\mathfrak{T}(b)).$$

Therefore

$$\mathfrak{T}(\mu(a, b)) = \mathfrak{T}(a)(\mathfrak{T}(b)),$$

i.e.,

$$\mathfrak{T}(\mu(x, y)) = \mathfrak{T}(x)(\mathfrak{T}(y)). \tag{2}$$

This equation represents a course-of-values recursion[51] for $\mathfrak{T}$. If we now define

$$\mathfrak{T}(e_k)[52] \equiv E_k,$$
$$\mathfrak{T}(a) \equiv I, \quad \text{if no term is associated with } a, \tag{3}$$

and define $\mathfrak{T}$ recursively by (2) if $\mu_3(x) = 0$, then $\mathfrak{T}(x)$ will be defined for all x and will have the property (1).

9. **The function** T. The constant T, mentioned in §4, is a formalization, so to speak, of the $\mathfrak{T}$ considered in 8.3. Its characteristic properties are the following:

$$a = \mathfrak{n}(\mathfrak{A}) \rightarrow T\mathcal{Z}_a = \mathfrak{A},$$
$$\vdash \mathcal{N}X \;\&\; \vdash \mathcal{N}Y \rightarrow T(MXY) = TX(TY),$$

which are motivated in 8.3. The preliminary analysis leading up to this function T is given in 9.1; the formal definition is given in 9.2; 9.3 and 9.4 are lemmas; while the final proof that the above two properties hold occupies 9.5 and 9.6.

9.1. *Preliminary analysis.* As pointed out in 8.3 the second property amounts to a course-of-values recursion for T, which, together with

$$T\mathcal{Z}_0 \equiv I, \qquad T\mathcal{Z}_{e_k} \equiv E_k, \tag{1}$$

essentially defines T. In order to take care of this recursion by the methods of §5 we must first consider a "course-of-values function," call it $\mathfrak{V}$, such that, intuitively,

---

[51] This expression is a translation of the German term 'Wertverlaufsrekursion' introduced by Péter (C 466.2). Cf. Hilbert-Bernays C 507.1, p. 326. The translation is used by Péter herself in Acta Szeged, vol. 9 (1940), p. 233.

[52] $e_k = \mathfrak{n}(E_k)$.

$$\mathfrak{V}\mathsf{Z}_x\mathsf{Z}_y = \begin{cases} T\mathsf{Z}_y & \text{if } y \leqq x, \\ I & \text{if } y > x. \end{cases} \tag{2}$$

If we can define such a $\mathfrak{V}$ then we can define T as $\lambda X.\mathfrak{V}XX$. This $\mathfrak{V}$ will be definable by a simple recursion of the kind considered in §5.

It is evidently in accordance with (1) to specify that

$$\mathfrak{V}\mathsf{Z}_0 = KI = \mathsf{Z}_0, \tag{3}$$

and with (2) to require that

$$\mathfrak{V}\mathsf{Z}_{n+1}\mathsf{Z}_m = \mathfrak{V}\mathsf{Z}_n\mathsf{Z}_m, \quad \text{if} \quad m \neq n+1,$$

while for $m = n+1$ we set

$$\mathfrak{V}\mathsf{Z}_{n+1}\mathsf{Z}_m = \mathfrak{U}_1\mathsf{Z}_{n+1},$$

where $\mathfrak{U}_1$ is yet to be defined. These can be combined by setting

$$\mathfrak{V}(\mathfrak{S}X) = \lambda Y.D_2(\mathfrak{U}_1(\mathfrak{S}X))(\mathfrak{V}XY)(\Delta_2(\mathfrak{S}X)Y). \tag{4}$$

Let us turn now to $\mathfrak{U}_1$. If we set

$$\mathfrak{U}_1 \equiv \lambda X.D_3(\mathfrak{U}_2X)(\mathfrak{U}_3X)(M_3X), \tag{5}$$

then it is evident that

$$\mathfrak{U}_1\mathsf{Z}_x = \begin{cases} \mathfrak{U}_2\mathsf{Z}_x & \text{if} \quad \mu_3(x) = 0, \\ \mathfrak{U}_3\mathsf{Z}_x & \text{if} \quad \mu_3(x) \neq 0. \end{cases}$$

Turning now to $\mathfrak{U}_2$ and setting

$$\mathfrak{U}_2 \equiv \lambda X.\mathfrak{V}X(M_1X)(\mathfrak{V}X(M_2X)),$$

we have that if $\mu_3(n) = 0$

$$\begin{aligned} T\mathsf{Z}_{n+1} &= \mathfrak{V}\mathsf{Z}_{n+1}\mathsf{Z}_{n+1} \\ &= \mathfrak{U}_1\mathsf{Z}_{n+1} \\ &= \mathfrak{U}_2\mathsf{Z}_{n+1} \\ &= \mathfrak{V}\mathsf{Z}_{n+1}(M_1\mathsf{Z}_{n+1})(\mathfrak{V}\mathsf{Z}_{n+1}(M_2\mathsf{Z}_{n+1})) \\ &= T(M_1\mathsf{Z}_{n+1})(T(M_2\mathsf{Z}_{n+1})) \qquad \text{[by (2)]}. \end{aligned}$$

Turn now to $\mathfrak{U}_3$, and define functions $\mathfrak{K}_0, \mathfrak{K}_1, \cdots, \mathfrak{K}_t$ by

$$\mathfrak{K}_i \equiv \lambda X.D_2E_i(\mathfrak{K}_{i+1}X)(\Delta_2\mathsf{Z}_{e_i}X) \qquad (i = 0, 1, \cdots, t-1),$$
$$\mathfrak{K}_j \equiv \lambda X.D_2E_iI(\Delta_2\mathsf{Z}_{e_i}X).$$

Then if we take $\mathfrak{K}_0$ for $\mathfrak{U}_3$ our definition of T is complete.

If we substitute for $\mathfrak{U}_1, \mathfrak{U}_2, \mathfrak{U}_3$ in (4) we have a recursive definition for $\mathfrak{V}$. This leads to the definition given formally in 9.2.

9.2. *Definitions.*

9.21. $\mathfrak{K}_i \equiv \lambda Y . D_2 E_i(\mathfrak{K}_{i+1} Y)(\Delta_2 Z_{e_i} Y)$ $(i = 0, 1, \cdots, t-1)$, and $\mathfrak{K}_t \equiv \lambda Y . D_2 E_t I(\Delta_2 Z_{e_t} Y)$.

9.22. $\mathfrak{U}_2 \equiv \lambda^3 VY . V(M_1 Y)(V(M_2 Y))$.

9.23. $\mathfrak{U}_1 \equiv \lambda^3 VY . D_2(\mathfrak{U}_2 VY)(\mathfrak{K}_0 Y)(M_3 Y)$.

9.24. $\mathfrak{G} \equiv \lambda^3 XVY . D_2(\mathfrak{U}_1 VY)(VY)(\Delta_2(\mathfrak{S} X) Y)$.

9.25. $\mathfrak{V} \equiv R\mathfrak{G}Z_0$.

9.26. $T \equiv \lambda X . \mathfrak{V}XX$.

*9.3. LEMMA. *If* $\vdash \mathcal{N}X$ & $\vdash \mathcal{N}Y$, *then*

$$\mathfrak{V}(X + Y)X = TX.$$

Proof. If $Y = Z_0$, then

$$\begin{aligned} \mathfrak{V}(X + Y)X &= \mathfrak{V}(X + Z_0)X \\ &= TX \qquad [9.26, 6.91]. \end{aligned}$$

Suppose the lemma true for a given Y[53]. Then

$$\begin{aligned} \mathfrak{V}(X + \mathfrak{S} Y)X &= \mathfrak{V}(\mathfrak{S}(X + Y))X && [6.91], \\ &= \mathfrak{G}(X + Y)(\mathfrak{V}(X + Y))X && [9.25; 5.53], \\ &= D_2\mathfrak{W}(\mathfrak{V}(X + Y)X)(\Delta_2(\mathfrak{S}(X + Y))X) && [9.24], \end{aligned}$$

where $\mathfrak{W}$ is a term which does not concern us for the moment. But, by 6.91,

$$\begin{aligned} \Delta_2(\mathfrak{S}(X + Y))X &= \Delta_2(X + \mathfrak{S} Y)X \\ &= ((X + \mathfrak{S} Y) \dot{-} X) + (X \dot{-} (X + \mathfrak{S} Y)) && [6.95], \\ &= \mathfrak{S} Y + Z_0 && [6.934, 6.935, 6.932], \\ &= \mathfrak{S} Y && [6.91]. \end{aligned}$$

Therefore

$$\begin{aligned} \mathfrak{V}(X + \mathfrak{S} Y)X &= \mathfrak{V}(X + Y)X && [5.32], \\ &= TX && \text{(by hypothesis of induction).} \end{aligned}$$

The lemma follows by 5.44.

*9.4. LEMMA. *If* $\vdash \mathcal{N}X$, *then* $T(\mathfrak{S} X) = D_2(\mathfrak{U}_2(\mathfrak{V}X)(\mathfrak{S} X))(\mathfrak{K}_0(\mathfrak{S} X))(M_3(\mathfrak{S} X))$.

Proof. We have

$$\begin{aligned} T(\mathfrak{S} X) &= \mathfrak{V}(\mathfrak{S} X)(\mathfrak{S} X) && [9.26], \\ &= \mathfrak{G}X(\mathfrak{V}X)(\mathfrak{S} X) && [9.25; 5.53], \\ &= D_2(\mathfrak{U}_1(\mathfrak{V}X)(\mathfrak{S} X))\mathfrak{W}(\Delta_2(\mathfrak{S} X)(\mathfrak{S} X)) && [9.24], \end{aligned} \tag{1}$$

(53) I.e., suppose $\vdash \mathcal{N}Y$ & $\vdash \mathfrak{M}Y$, where

$$\mathfrak{M} \equiv \lambda Y . Q(\mathfrak{V}(X + Y)X)(TX).$$

where $\mathfrak{W}$ is a term which does not concern us. By 6.95, $\Delta_2(\mathfrak{S}X)(\mathfrak{S}X) = \mathfrak{Z}_0$. Therefore [(1); 5.31]

$$T(\mathfrak{S}X) = \mathfrak{U}_1(\mathfrak{V}X)(\mathfrak{S}X).$$

The lemma follows by 9.23.

*9.5. THEOREM. *If* $\vdash \mathcal{N}X$ & $\vdash \mathcal{N}Y$, *then* $T(MXY) = TX(TY)$.

Proof. By 3.522, $|\mu(x, y)| = 0$. Therefore [6.944] $\mu(x, y) = \sigma(\delta_1(\mu(x, y)))$. Therefore [7.51, hypothesis]

(1) $$MXY = \mathfrak{S}\mathfrak{A},$$

where $\mathfrak{A} \equiv \Delta_1(MXY)$. Therefore

$$\begin{aligned} T(MXY) &= T(\mathfrak{S}\mathfrak{A}) \\ &= D_2(\mathfrak{U}_2(\mathfrak{V}\mathfrak{A})(\mathfrak{S}\mathfrak{A}))(\mathfrak{K}_0(\mathfrak{S}\mathfrak{A}))(M_3(\mathfrak{S}\mathfrak{A})) && [9.4], \\ &= \mathfrak{U}_2(\mathfrak{V}\mathfrak{A})(\mathfrak{S}\mathfrak{A}) && [(1), 8.23, 5.31], \\ &= \mathfrak{V}\mathfrak{A}(M_1(\mathfrak{S}\mathfrak{A}))(\mathfrak{V}\mathfrak{A}(M_2(\mathfrak{S}\mathfrak{A}))) && [9.22], \\ (2) \qquad &= \mathfrak{V}\mathfrak{A}X(\mathfrak{V}\mathfrak{A}Y) && [(1), 8.21]. \end{aligned}$$

But, by 3.522, $\mu(x, y) > x$ & $\mu(x, y) > y$. Therefore

$$\delta_1(\mu(x, y)) \geqq x.$$

Hence

$$\delta_1(\mu(x, y)) = x + (\delta_1(\mu(x, y)) \dot{-} x).$$

It follows [7.51] that $\mathfrak{A} = X + (\mathfrak{A} \dot{-} X)$. Similarly $\mathfrak{A} = Y + (\mathfrak{A} \dot{-} Y)$. Hence, by 9.3

$$\mathfrak{V}\mathfrak{A}X = TX, \qquad \mathfrak{V}\mathfrak{A}Y = TY.$$

Substituting these in (2) we have the theorem.

9.6. THEOREM. *If* $a = \mathfrak{n}(\mathfrak{A})$, *then* $\mathfrak{A} = T\mathfrak{Z}_a$.

Proof. We proceed by induction on the construction of $\mathfrak{A}$. In 9.61 we take the case where $\mathfrak{A}$ is a primitive term, in 9.62 the induction from $\mathfrak{B}$ and $\mathfrak{C}$ to $\mathfrak{A} = \mathfrak{B}\mathfrak{C}$.

9.61. Suppose $\mathfrak{A} \equiv E_k$; then $a = \mathfrak{n}(\mathfrak{A}) = e_k > 0$. Let

$$b = e_k - 1.$$

Then

$$a = b + 1.$$

Therefore [3.44]

(1) $$\mathfrak{Z}_a = \mathfrak{S}\,\mathfrak{Z}_b;$$

hence [9.4]

(2) $$T\mathfrak{Z}_a = D_2\mathfrak{W}(\mathfrak{K}_0\mathfrak{Z}_a)(M_2\mathfrak{Z}_a),$$

where $\mathfrak{W}$ is a term which does not concern us. By 8.1, 3.524, $\mu_3(a)>0$. Hence there is a number c such that $\mu_3(a)=c+1$. Hence [7.54, 3.44] $M_2\mathfrak{Z}_a=\mathfrak{S}\,\mathfrak{Z}_c$. Therefore [(2); 5.32]

(3) $$T\mathfrak{Z}_a = \mathfrak{K}_0\mathfrak{Z}_a.$$

Now for $i<k$, $\delta_2(e_i, a)>0$. Therefore [7.6]

$$\Delta_2\mathfrak{Z}_{e_i}\mathfrak{Z}_a = \mathfrak{S}\,(\Delta_1(\Delta_2\mathfrak{Z}_{e_i}\mathfrak{Z}_a)).$$

Thus [9.21; 5.32]

(4) $$\mathfrak{K}_i\mathfrak{Z}_a = \mathfrak{K}_{i+1}\mathfrak{Z}_a.$$

But since $\Delta_2\mathfrak{Z}_{e_k}\mathfrak{Z}_a=\mathfrak{Z}_0$, we have [5.31]

(5) $$\mathfrak{K}_k\mathfrak{Z}_a = E_k.$$

By (3), (4), (5) we have $T\mathfrak{Z}_a=\mathfrak{A}$.

9.62. Suppose now $\mathfrak{A}\equiv\mathfrak{B}\mathfrak{C}$, and that the theorem is true for $\mathfrak{B}$ and $\mathfrak{C}$. Let $b\equiv\mathfrak{n}(\mathfrak{B})$ and $c\equiv\mathfrak{n}(\mathfrak{C})$; then [3.521] $a=\mu(b, c)$. Hence [7.54]

$$\mathfrak{Z}_a = M\mathfrak{Z}_b\mathfrak{Z}_c.$$

Therefore

$$\begin{aligned} T\mathfrak{Z}_a &= T(M\mathfrak{Z}_b\mathfrak{Z}_c) \\ &= T\mathfrak{Z}_b(T\mathfrak{Z}_c) && [9.5], \\ &= \mathfrak{B}\mathfrak{C} && [\text{hypothesis of induction}], \\ &= \mathfrak{A}, \end{aligned}$$

q.e.d.

PART IV. FORMAL ENUMERATION OF THEOREMS

The object of this part is to define a term Θ which gives a formal enumeration of the theorems of the system, in the sense that every term of the form $T(\Theta\mathfrak{Z}_n)$ is assertible, and vice versa. Θ has then the property that

$$\vdash \mathfrak{N}X \rightarrow_X T(\Theta X).$$

Θ is the formalization of a primitive recursive function $\theta(x)$; this $\theta(x)$ and some numerical functions related to it are defined in §10. The remaining sections concern the formalizations of these various functions. The final result is established in §14.

10. The function θ and related functions.

10.1. *Enumeration of the axioms.* We define first a function $\alpha(x)$ which enumerates the numbers of the axioms. By hypothesis there are only a finite number of axioms; let these be $A_0, A_1, \cdots, A_s$, and let $a_0, a_1, \cdots, a_s$ respectively be their associated numbers. Then let

$$\alpha_1(x) \equiv \sum_{k=0}^{x} \left| \delta_2(x, k) \right| \cdot a_k. \tag{1}$$

We can easily define a function $\beta(x)$ whose value for every x is one of the a_i, and which takes each of the a_i at least once; for instance let $\beta(x)$ be a primitive recursive function which takes on only the values $0, 1, \cdots, s$ (e.g.,

$$x \dot{-} \left[\frac{x}{s+1} \right] \cdot (s+1)).$$

Then set

$$\alpha(x) \equiv \alpha_1(x) + \|x \dot{-} s\| \cdot \alpha_1(\beta(x)),$$

$$= \sum_{k=0}^{x} \left| \delta_2(x, k) \right| \cdot \alpha_1(k) + \|x \dot{-} s\| \cdot \alpha_1(\beta(x)). \tag{2}$$

10.2. *Functions connected with the rules.* Suppose we have a rule of the form

$$\vdash \mathfrak{X}_1 \,\&\, \cdots \,\&\, \vdash \mathfrak{X}_n \rightarrow_X \mathfrak{X}_{n+1},$$

where the $\mathfrak{X}_j$ are combinations of constants and the indeterminates $X_1, \cdots, X_m$. We must furthemore assume that any indeterminate which actually occurs in $\mathfrak{X}_{n+1}$ actually occurs in some $\mathfrak{X}_j$ for $j<n$. We seek to find a recursive function $\rho(x_1, \cdots, x_n)$, such that if $\mathfrak{X}_1^*, \cdots, \mathfrak{X}_{n+1}^*$ are terms obtained by substituting constants $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ for $X_1, \cdots, X_m$ in $\mathfrak{X}_1, \cdots, \mathfrak{X}_{n+1}$ respectively, and if $b_j = \mathfrak{n}(\mathfrak{X}_j^*)$, then $\rho(b_1, \cdots, b_n) = b_{n+1}$; furthermore ρ is to have the property that if $c_1, \cdots, c_n$ are numbers associated with theorems (whether obtained by substitution from the $\mathfrak{X}_j$ or not) then $\rho(c_1, \cdots, c_n)$ is also. In order to obtain such a ρ we must first define functions $\phi(\mathfrak{X}_j; x_1, \cdots, x_m)$ expressing the above b_j in terms of $\mathfrak{n}(\mathfrak{A}_1), \cdots, \mathfrak{n}(\mathfrak{A}_m)$ (this is done in 10.21), and then (in 10.22) functions allowing, so to speak, the inversion of the ϕ's. The definition of ρ is then given in 10.23.

10.21. Let $\mathfrak{X}$ be any combination of constants and the indeterminates $X_1, \cdots, X_m$. With each such $\mathfrak{X}$ we associate a function $\phi(\mathfrak{X}; x_1, \cdots, x_m)$ as specified, by intuitive recursion on the construction of $\mathfrak{X}$, in 10.211–10.213. The properties 10.214 and 10.215 then follow by induction.

10.211. If $\mathfrak{C}$ is a constant(54), then

$$\phi(\mathfrak{C}; x_1, \cdots, x_m) \equiv \mathfrak{n}(\mathfrak{C}).$$

(54) It is sufficient to make this definition for $\mathfrak{C}$ a primitive constant; it then follows by 10.213 for any constant.

10.212. For $i=1, 2, \cdots, m$,

$$\phi(X_i; x_1, \cdots, x_m) \equiv x_i.$$

10.213. $\phi(\mathfrak{YB}; x_1, \cdots, x_m) \equiv \mu(\phi(\mathfrak{Y}; x_1, \cdots, x_m), \phi(\mathfrak{B}; x_1, \cdots, x_m))$.

10.214. If $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ are any constants and if $a_i = \mathfrak{n}(\mathfrak{A}_i)$, then

$$\phi(\mathfrak{X}; a_1, \cdots, a_m) = \mathfrak{n}\left(\begin{bmatrix} \mathfrak{A}_1, \cdots, \mathfrak{A}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}\right).$$

10.215. If $X_{j_1}, X_{j_2}, \cdots, X_{j_p}$ are the indeterminates actually occurring in $\mathfrak{X}$, then, if

$$\phi(\mathfrak{X}; a_1, \cdots, a_m) = \phi(\mathfrak{X}; b_1, \cdots, b_m),$$

we have

$$a_{j_1} = b_{j_1} \,\&\, \cdots \,\&\, a_{j_p} = b_{j_p};$$

furthermore $\phi(\mathfrak{X}; a_1, \cdots, a_m)$ is not the number of any constant unless $a_{j_1}, \cdots, a_{j_p}$ are also [3.52]. The first statement may, moreover, be formally established in the form

$$(1) \qquad \begin{aligned} &\left| \delta_2(\phi(\mathfrak{X}; x_1, \cdots, x_m), \phi(\mathfrak{X}; y_1, \cdots, y_m)) \right| \\ &\qquad = \left| \delta_2(x_{j_1}, y_{j_1}) + \cdots + \delta_2(x_{j_p}, y_{j_p}) \right|. \end{aligned}$$

10.22. Consider now a rule of the form stated in the introduction to 10.2. We shall suppose that the $X_1, \cdots, X_m$ are all the indeterminates which actually occur in the $\mathfrak{X}_1, \cdots, \mathfrak{X}_n$; the rule is then such that these include all which actually occur in $\mathfrak{X}_{n+1}$ also. We define now functions ψ_i, such that, if $b_1, \cdots, b_n$ are the numbers of terms arising by substitution of $\mathfrak{A}_1, \cdots, \mathfrak{A}_m$ for $X_1, \cdots, X_m$ in the $\mathfrak{X}_j$, then $\psi_i(b_1, \cdots, b_n) = \mathfrak{n}(\mathfrak{A}_i)$.

In making these definitions it is convenient to abbreviate as follows:

$$\begin{aligned} \phi_j(x_1, \cdots, x_m) &\equiv \phi(\mathfrak{X}_j; x_1, \cdots, x_m), \\ \Sigma y &\equiv y_1 + y_2 + \cdots + y_n, \\ \psi_i^* &\equiv \psi_i(y_1, \cdots, y_n), \\ \mathfrak{m}_i &= \sum_{j=1}^{n} \delta_2(y_j, \phi_j(\psi_1^*, \cdots, \psi_{i-1}^*, x_i, \cdots, x_m)). \end{aligned}$$

Then the definitions are

$$(1) \qquad \begin{aligned} \psi_0^* &\equiv (\exists x_1 \leqq \Sigma y)(\exists x_2 \leqq \Sigma y) \cdots (\exists x_m \leqq \Sigma y)\mathfrak{m}_1, \\ \psi_i^* &\equiv (\epsilon x_i \leqq \Sigma y)(\exists x_{i+1} \leqq \Sigma y) \cdots (\exists x_m \leqq \Sigma y)\mathfrak{m}_i \\ &\qquad\qquad (i = 1, 2, \cdots, m). \end{aligned}$$

From these definitions it follows that if $\psi_0(y_1, \cdots, y_n) = 0$, then $\phi_j(\psi_1^*, \cdots, \psi_m^*) = y_j$; i.e.,

$$(2) \qquad |\psi_0(y_1, \cdots, y_n)| \cdot \delta_2(y_j, \phi_j(\psi_1^*, \cdots, \psi_m^*)) = 0.$$

10.23. In terms of the function defined in 10.21 and 10.22 the definition of ρ is the following:

$$\rho(y_1, \cdots, y_n) \equiv |\psi_0^*| \cdot \phi_{n+1}(\psi_1^*, \cdots, \psi_m^*) + \|\psi_0^*\| \cdot \mathfrak{n}(A_0).$$

This ρ then has the properties stated in the introduction to 10.2.

10.3 *Definition of* θ. The definition of $\theta(x)$, which enumerates with repetitions the numbers associated with the theorems, is essentially due to Kleene([55]). The idea of this enumeration is as follows: First we divide the theorems into stages. In the 0th stage we put only the axiom A_0; and in the $(k+1)$st stage we put the axiom A_k (i.e., $\mathfrak{T}(\alpha(k))$) together with the theorems obtained by applying the rules in all possible ways to the theorems in all the preceding stages. We may suppose that all the rules have the same number of premises; for if any rule has less than the maximum number, we increase the number by repeating premises. Then within each stage we can arrange the theorems, first according to the rule used, then according to the first premise, then according to the second premise, etc. In such a scheme every theorem will have at least one number, and this number can be effectively ascertained when once a proof of the theorem is known.

This numbering of the theorems is determined by a set of recursive functions as follows([56]):

$\zeta(k)$ is the number of theorems in the stages preceding the kth.

$\gamma(x)$ = the stage of the theorem numbered x.

$\nu(x)$ = the number of theorem x in its stage.

$\pi_0(x)$ = the number of the rule used([57]).

$\pi_k(x)$ = the number of the kth premise $(k = 1, 2, \cdots, n)$.

$\xi_k(x)$ = the number of the theorem in the group in the same stage having the same rule and (for $k > 0$) the same k first premises.

Let us suppose there are r rules of n premises each. Let the rules be numbered 0, 1, 2, $\cdots$, $r-1$; and let ρ_k be the function associated with the kth rule as in 10.2. Then the above functions have the following recursive definitions:

$$(1) \qquad \begin{aligned} \zeta(0) &= 0, \\ \zeta(x+1) &= \zeta(x) + 1 + r \cdot (\zeta(x))^n, \end{aligned}$$

([55]) C 497.2.

([56]) In the enumeration defined below these theorems are placed first and A_k last, and the numbering in each of the substages begins with 0.

([57]) If the nth theorem is an axiom, $\pi_0(n) = r$, where r (as below) is the number of rules.

$$\gamma(x) \equiv (ez \leqq x)[(\zeta(z) \dot{-} x) + ((x+1) \dot{-} \zeta(z+1))]. \tag{2}$$

$$\nu(x) \equiv x \dot{-} \zeta(\gamma(x)), \tag{3}$$

$$\pi_0(x) \equiv \left[\frac{\nu(x)}{(\zeta(\gamma(x)))^n}\right], \tag{4}$$

$$\xi_0(x) \equiv \nu(x) \dot{-} \pi_0(x) \cdot (\zeta(\gamma(x)))^n, \tag{5}$$

$$\pi_{k+1}(x) \equiv \left[\frac{\xi_k(x)}{(\zeta(\gamma(x)))^{n-k-1}}\right] \qquad (k = 0, 1, \cdots, n-1), \tag{6}$$

$$\xi_{k+1}(x) \equiv \xi_k(x) \dot{-} \pi_{k+1}(x) \cdot (\zeta(\gamma(x)))^{n-k-1} \qquad (k = 0, 1, \cdots, n-2). \tag{7}$$

Then the recursive equations for $\theta(x)$ are

$$\theta(0) = \alpha(0); \tag{8}$$

and, for $x>0$,

$$\theta(x) = \sum_{k=0}^{r-1} \left| \delta_2(\pi_0(x), k) \right| \cdot \rho_k(\vartheta(\pi_1(x)), \cdots, \theta(\pi_n(x))) + \left\| \pi_0(x) \dot{-} (r-1) \right\| \cdot \alpha(\gamma(x)). \tag{9}$$

Since for $x>0$, $i>0$,

$$\pi_i(x) \leqq \nu(x) < x, \tag{10}$$

this is a course-of-values recursion in the sense of R. Péter[58] for θ. According to her work it can be replaced by a primitive recursion; i.e., there exists a primitive recursive function having the properties (8) and (9).

11. **Proof by cases.** In the foregoing we have frequently had to deal with functions of the form of 6.98. In proving theorems concerning the formalizations of such functions we shall need a technique for obtaining proofs by cases—i.e., proof of a theorem by considering separately the case where some numerical term is $\mathfrak{Z}_0$ or is not, etc. The necessary theorems in this connection are developed in the present section.

11.1. THEOREM. *If* $\mathfrak{A}, \mathfrak{B}, \mathfrak{M}$ *are any terms such that*

(a) $\vdash \mathfrak{M} \rightarrow \vdash \mathcal{N}\mathfrak{A}$,

(b) $\vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathfrak{Z}_0 \rightarrow \vdash \mathfrak{B}$,

(c) $\vdash \mathfrak{M} \,\&\, |\mathfrak{A}| = \mathfrak{Z}_0 \rightarrow \vdash \mathfrak{B}$,

then

$$\vdash \mathfrak{M} \rightarrow \vdash \mathfrak{B}.$$

Proof. Let $\mathfrak{L} \equiv \lambda X . \Lambda \mathfrak{M}(Q\mathfrak{A}X)$. Then [5.42]

---

(58) See second footnote to 8.3.

(1) $$\vdash \mathfrak{L}Y \rightleftarrows_Y \vdash \mathfrak{M} \,\&\, \mathfrak{A} = Y.$$

Therefore

$$\vdash \mathfrak{L}\mathfrak{Z}_0 \rightarrow \vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathfrak{Z}_0 \rightarrow \vdash \mathfrak{B} \qquad [(b)].$$

Hence [3.43]

$$\vdash K(\mathfrak{L}\mathfrak{Z}_0)X \rightarrow_X \vdash K\mathfrak{B}(IX).$$

Thus [3.65]

(2) $$\vdash F(K(\mathfrak{L}\mathfrak{Z}_0))(K\mathfrak{B})I.$$

Again, by (1),

$$\vdash \mathcal{N}Y \,\&\, \vdash \mathfrak{L}(\mathfrak{S}\,Y) \rightarrow_Y \vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathfrak{S}\,Y \,\&\, \vdash \mathcal{N}Y \rightarrow_Y \vdash \mathfrak{M} \,\&\, |\mathfrak{A}| = \mathfrak{Z}_0 \qquad [6.941],$$
$$\rightarrow_Y \vdash \mathfrak{B} \qquad [(c)].$$

Hence [3.43, 3.41]

$$\vdash \mathcal{N}Y \,\&\, \vdash K(\mathfrak{L}(\mathfrak{S}\,Y))X \rightarrow_{X,Y} \vdash K\mathfrak{B}(IX).$$

Therefore [3.65]

(3) $$\vdash \mathcal{N}Y \rightarrow_Y \vdash F(K(\mathfrak{L}(\mathfrak{S}\,Y)))(K\mathfrak{B})I.$$

Then the following can be successively deduced:

$$\vdash \mathcal{N}Y \rightarrow_Y \vdash F(K(\mathfrak{L}Y))(K\mathfrak{B})I \qquad [(2), (3), 5.44],$$
$$\vdash \mathcal{N}Y \,\&\, \vdash K(\mathfrak{L}Y)X \rightarrow_{Y,X} \vdash K\mathfrak{B}(IX) \qquad [3.63],$$
$$\vdash \mathcal{N}Y \,\&\, \vdash \mathfrak{L}Y \rightarrow_Y \vdash \mathfrak{B} \qquad [3.43],$$
$$\vdash \mathcal{N}Y \,\&\, \vdash \mathfrak{M} \,\&\, \mathfrak{A} = Y \rightarrow_Y \vdash \mathfrak{B} \qquad [(1)],$$
$$\vdash \mathcal{N}\mathfrak{A} \,\&\, \vdash \mathfrak{M} \rightarrow \vdash \mathfrak{B} \qquad [3.2].$$

The hypothesis $\vdash \mathcal{N}\mathfrak{A}$ is superfluous by (a), and the theorem is proved.

11.2. THEOREM. *If* $\mathfrak{A}, \mathfrak{B}, \mathfrak{M}$ *are terms such that*

(a) $\vdash \mathfrak{M} \rightarrow \vdash \mathcal{N}\mathfrak{A}$,

and if for $k = 0, 1, 2, \cdots, p$ *we have*

(b) $\vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathfrak{Z}_k \rightarrow \vdash \mathfrak{B}$,

then

$$\vdash \mathfrak{M} \,\&\, \mathfrak{A} \dot{-} \mathfrak{Z}_p = \mathfrak{Z}_0 \rightarrow \vdash \mathfrak{B}.$$

Proof. For $p = 0$ the conclusion is the same as the hypothesis [by 6.93]. Suppose the theorem true for $p = q$; I shall show that it then follows for $p = q+1$.

Suppose then $p=q+1$ and the theorem holds for q. Then the first $q+1$ hypotheses (b) (viz., for $k=0, 1, \cdots, q$) imply that

$$\vdash \mathfrak{M} \,\&\, \mathfrak{A} \doteq \mathfrak{Z}_q = \mathfrak{Z}_0 \rightarrow \vdash \mathfrak{B}. \tag{1}$$

Let $\mathfrak{M}_1 \equiv \Lambda\mathfrak{M}(Q(\mathfrak{A} \doteq \mathfrak{Z}_p)\mathfrak{Z}_0)$. Then [5.42]

$$\vdash \mathfrak{M}_1 \rightleftarrows \vdash \mathfrak{M} \,\&\, \mathfrak{A} \doteq \mathfrak{Z}_p = \mathfrak{Z}_0. \tag{2}$$

Hence

$$\begin{aligned} \vdash \mathfrak{M}_1 \,\&\, \mathfrak{A} \doteq \mathfrak{Z}_q = \mathfrak{Z}_0 &\rightarrow \vdash \mathfrak{M} \,\&\, \mathfrak{A} \doteq \mathfrak{Z}_q = \mathfrak{Z}_0 \\ &\rightarrow \vdash \mathfrak{B} \qquad [(1)]. \end{aligned} \tag{3}$$

On the other hand, by 6.86, the equation

$$|x \doteq \sigma(y)| \cdot \|x \doteq y\| \cdot \delta_2(x, \sigma(y)) = 0$$

is valid in $\mathfrak{r}$; hence, by 7.51, we know that

$$\vdash \mathfrak{N}X \,\&\, \vdash \mathfrak{N}Y \rightarrow_{X,Y} |X \doteq \mathfrak{S}Y| \cdot \|X \doteq Y\| \cdot \Delta_2 X(\mathfrak{S}Y) = \mathfrak{Z}_0.$$

In this we can substitute [2.61] $\mathfrak{A}$ for X and $\mathfrak{Z}_q$ for Y; the result, by (a), 5.27, 3.44, shows that

$$|\mathfrak{A} \doteq \mathfrak{Z}_p| \cdot \|\mathfrak{A} \doteq \mathfrak{Z}_q\| \cdot \Delta_2 \mathfrak{A} \mathfrak{Z}_p = \mathfrak{Z}_0.$$

Therefore [(2), 7.63]

$$\vdash \mathfrak{M}_1 \rightarrow \|\mathfrak{A} \doteq \mathfrak{Z}_q\| \cdot \Delta_2 \mathfrak{A} \mathfrak{Z}_p = \mathfrak{Z}_0.$$

Hence

$$\begin{aligned} \vdash \mathfrak{M}_1 \,\&\, |\mathfrak{A} \doteq \mathfrak{Z}_q| = \mathfrak{Z}_0 &\rightarrow \Delta_2 \mathfrak{A} \mathfrak{Z}_p = \mathfrak{Z}_0 \qquad [7.63], \\ &\rightarrow \mathfrak{A} = \mathfrak{Z}_p \qquad [7.62]. \end{aligned} \tag{4}$$

Thus

$$\begin{aligned} \vdash \mathfrak{M}_1 \,\&\, |\mathfrak{A} \doteq \mathfrak{Z}_q| = \mathfrak{Z}_0 &\rightarrow \vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathfrak{Z}_p \qquad [(2), (4)], \\ &\rightarrow \vdash \mathfrak{B} \qquad [(b)], \end{aligned} \tag{5}$$

and thus [(3), (5), 11.1]

$$\vdash \mathfrak{M}_1 \rightarrow \vdash \mathfrak{B}.$$

The conclusion of the theorem follows by (2).

11.3. THEOREM. *If $\mathfrak{A}, \mathfrak{B}, \mathfrak{M}$, are terms such that*
(a) $\vdash \mathfrak{M} \rightarrow \vdash \mathfrak{N}\mathfrak{A}$,
(b) *for* $k=0, 1, \cdots, p$, $\vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathfrak{Z}_k \rightarrow \vdash \mathfrak{B}$,
(c) $\vdash \mathfrak{M} \,\&\, |\mathfrak{A} \doteq \mathfrak{Z}_p| = \mathfrak{Z}_0 \rightarrow \vdash \mathfrak{B}$,
then

$$\vdash \mathfrak{M} \rightarrow \vdash \mathfrak{B}.$$

Proof. This follows by 11.2 and 11.1.

11.4. THEOREM. *If* $\mathfrak{A}, \mathfrak{B}_0, \cdots, \mathfrak{B}_p, \mathfrak{C}, \mathfrak{M}$ *are* $(X_1, \cdots, X_m)$*-terms, of which* $\mathfrak{A}, \mathfrak{B}_0, \cdots, \mathfrak{B}_p, \mathfrak{C}$ *are formalizations of numerical expressions, and* $\mathfrak{Q}$ *is a constant such that*

(a) $\vdash \mathfrak{M} \rightarrow_X \vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m$,

(b) $\vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathcal{Z}_k \rightarrow_X \vdash \mathfrak{Q}\mathfrak{B}_k$ $(k = 0, 1, 2, \cdots, p)$,

(c) $\vdash \mathfrak{M} \,\&\, |\mathfrak{A} \dot{-} \mathcal{Z}_p| = \mathcal{Z}_0 \rightarrow_X \vdash \mathfrak{Q}\mathfrak{C}$;

and if $\mathfrak{G}$ *is defined by*

$$\mathfrak{G} \equiv \lambda^m X_1 \cdots X_m . \left[\sum_{k=0}^{p} |\Delta_2 \mathfrak{A} \mathcal{Z}_k| \cdot \mathfrak{B}_k + \|\mathfrak{A} \dot{-} \mathcal{Z}_p\| \cdot \mathfrak{C} \right];$$

then

$$\vdash \mathfrak{M} \rightarrow_X \vdash \mathfrak{Q}(\mathfrak{G} X_1 \cdots X_m).$$

Proof[59]. Since $\mathfrak{G}$ is the formalization of the $\mathfrak{g}$ of 6.98 [by 7.52], we have [by 7.51, 7.22]

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow |\Delta_2 \mathfrak{A} \mathcal{Z}_k| \cdot \Delta_2(\mathfrak{G} X_1 \cdots X_m)\mathfrak{B}_k = \mathcal{Z}_0.$$

Therefore [(a); 7.62; 7.63]

$$\vdash \mathfrak{M} \,\&\, \mathfrak{A} = \mathcal{Z}_k \rightarrow \mathfrak{G} X_1 \cdots X_m = \mathfrak{B}_k. \tag{1}$$

Similarly,

$$\vdash \mathcal{N}X_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}X_m \rightarrow \|\mathfrak{A} \dot{-} \mathcal{Z}_p\| \cdot \Delta_2(\mathfrak{G} X_1 \cdots X_m)\mathfrak{C} = \mathcal{Z}_0.$$

Hence [(a); 7.63]

$$\vdash \mathfrak{M} \,\&\, |\mathfrak{A} \dot{-} \mathcal{Z}_p| = \mathcal{Z}_0 \rightarrow \mathfrak{G} X_1 \cdots X_m = \mathfrak{C}. \tag{2}$$

The theorem now follows by 11.3, 7.22.

12. **Formalization of the axiom enumeration.**

THEOREM. *If* $\mathcal{A} \equiv \mathfrak{F}(\alpha)$, *then* $\vdash \mathcal{N}X \rightarrow_X T(\mathcal{A}X)$.

Proof. Let $\mathcal{A}_1 \equiv \mathfrak{F}(\alpha_1)$. Then, since for $k = 0, 1, \cdots, s$,

$$T(\mathcal{A}_1 \mathcal{Z}_k) = T\mathcal{Z}_{\alpha_1(k)} \qquad [7.54],$$

---

[59] The referee has pointed out that the proof as given is not valid for the case that $\mathfrak{A}, \mathfrak{B}_0, \cdots, \mathfrak{B}_p, \mathfrak{C}$ are unrestricted $(X_1, \cdots, X_m)$-terms such that

$$\vdash \mathfrak{M} \rightarrow_X \vdash \mathcal{N}\mathfrak{A} \,\&\, \vdash \mathcal{N}\mathfrak{B}_0 \cdots \vdash \mathcal{N}\mathfrak{B}_p \,\&\, \vdash \mathcal{N}\mathfrak{C};$$

but that the theorem may easily be extended to cover that case—in fact, the theorem may first be established for the case that $\mathfrak{A}, \mathfrak{B}_0, \cdots, \mathfrak{B}_p, \mathfrak{C}$ are indeterminates and the general case derived by substitution. However, this more general theorem is not needed in this paper.

$$T(\mathsf{Z}_{\alpha_1(k)}) = \mathfrak{T}(\alpha_1(k)) \qquad [9.6],$$
$$= A_k \qquad [\text{definition of } \alpha_1],$$

we have

$$(1) \qquad \vdash T(\mathcal{A}_1\mathsf{Z}_k) \qquad (k = 0, 1, \cdots, s).$$

Therefore [11.2]

$$(2) \qquad \vdash \mathcal{N}X \,\&\, X \dot{-} \mathsf{Z}_s = \mathsf{Z}_0 \rightarrow \vdash T(\mathcal{A}_1X).$$

Next, let $\mathfrak{B} \equiv \mathfrak{F}(\beta)$. Then, since $\beta(x) \dot{-} s = 0$, we have [7.51] $\vdash \mathcal{N}X \rightarrow \mathfrak{B}X \dot{-} \mathsf{Z}_s = \mathsf{Z}_0$; also [7.21, 3.63] $\vdash \mathcal{N}X \rightarrow \vdash \mathcal{N}(\mathfrak{B}X)$. Hence by substitution in (2) we have

$$(3) \qquad \vdash \mathcal{N}X \rightarrow_X \vdash T(\mathcal{A}_1(\mathfrak{B}X)).$$

Now by (2) of 10.1 and 7.52,

$$\mathcal{A} \equiv \lambda X. \sum_{k=0}^{s} \left| \Delta_2 X\mathsf{Z}_k \right| \cdot \mathcal{A}_1\mathsf{Z}_k + \left\| X \dot{-} \mathsf{Z}_s \right\| \cdot \mathcal{A}_1(\mathfrak{B}X).$$

Hence the theorem follows by 11.4 from (1) and (3).

Remark. This theorem is the only place in the paper where the finiteness of the axiom system plays an essential role.

13. **Formalization of ρ and related functions.** Let $\mathfrak{R}$ be the formalization of the ρ defined in 10.2. We are concerned in this section with the proof that

$$\vdash \mathcal{N}Y_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}Y_n \,\&\, \vdash TY_1 \,\&\, \cdots \,\&\, \vdash TY_n \rightarrow \vdash T(\mathfrak{R}Y_1 \cdots Y_n),$$

under the hypothesis that the rule with which ρ is associated in 10.2 is one of the rules of procedure.

In this section the formalizations of the various functions considered in 10.2 are symbolized as follows:

$$\Phi(\mathfrak{X}) \equiv \lambda^m X_1 \cdots X_m. \mathfrak{F}(\phi(\mathfrak{X}; x_1, \cdots, x_m)),$$
$$\Phi_j \equiv \mathfrak{F}(\phi_j) \qquad (j = 1, 2, \cdots, n),$$
$$\Psi_i \equiv \mathfrak{F}(\psi_i) \qquad (i = 0, 1, 2, \cdots, m).$$

Let us also define

$$\mathfrak{M} \equiv \Lambda_{2n}(\mathcal{N}Y_1) \cdots (\mathcal{N}Y_n)(TY_1) \cdots (TY_n),$$
$$\mathfrak{Y}_i \equiv \Psi_i Y_1 \cdots Y_n \qquad (i = 0, 1, 2, \cdots, m).$$

The course of the proof is as follows: In 13.1 a lemma concerning $\Phi(\mathfrak{X})$ is established. Then in 13.2 it is shown that

$$\vdash \mathfrak{M} \,\&\, \mathfrak{Y}_0 = \mathsf{Z}_0 \rightarrow_Y \vdash T(\mathfrak{R}Y_1 \cdots Y_n),$$

while in 13.3 we shall see that

$$\vdash \mathfrak{M} \,\&\, |\, \mathfrak{Y}_0| = \mathfrak{Z}_0 \rightarrow_Y \vdash T(\mathfrak{R} Y_1 \cdots Y_n).$$

The theorem concerning $\mathfrak{R}$ then follows by 11.1 and 5.43.

*13.1. LEMMA. *On the assumption that*

$$\vdash \mathfrak{N}\mathfrak{U}_1 \,\&\, \cdots \,\&\, \vdash \mathfrak{N}\mathfrak{U}_m,$$

it follows that

$$T(\Phi(\mathfrak{X})\mathfrak{U}_1 \cdots \mathfrak{U}_m) = \begin{bmatrix} T\mathfrak{U}_1, \cdots, T\mathfrak{U}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}.$$

Proof. Let

$$\mathfrak{X}' = \Phi(\mathfrak{X})\mathfrak{U}_1 \cdots \mathfrak{U}_m,$$

$$\mathfrak{X}^* = \begin{bmatrix} T\mathfrak{U}_1, \cdots, T\mathfrak{U}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}.$$

Then, by 3.32, it follows that

$$\mathfrak{X}' = \begin{bmatrix} \mathfrak{U}_1, \cdots, \mathfrak{U}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{F}(\phi(\mathfrak{X}; x_1, \cdots, x_m)). \tag{1}$$

To prove the lemma we proceed by intuitive induction on the construction of $\mathfrak{X}$.

1. Suppose $\mathfrak{X} \equiv \mathfrak{C}$, $c \equiv \mathfrak{n}(\mathfrak{C})$, where $\mathfrak{C}$ is a constant. Then [(1), 10.211, 7.53]

$$\mathfrak{X}' = \mathfrak{Z}_c.$$

Therefore [9.6]

$$T\mathfrak{X}' = \mathfrak{C} = \mathfrak{X}^* \qquad [2.6].$$

2. Suppose $\mathfrak{X} \equiv X_i$. Then [(1), 10.212, 7.12]

$$\mathfrak{X}' = \mathfrak{U}_i.$$

Therefore [2.6]

$$T\mathfrak{X}' = T\mathfrak{U}_i = \mathfrak{X}^*.$$

3. Suppose, finally, that $\mathfrak{X} \equiv \mathfrak{X}_1\mathfrak{X}_2$ and that the theorem holds for $\mathfrak{X}_1$ and $\mathfrak{X}_2$. Then [(1), 7.52, 10.213]

$$\mathfrak{X}' = M\mathfrak{X}_1'\mathfrak{X}_2'. \tag{2}$$

But [(1), 7.22, hypothesis]

$$\vdash \mathfrak{N}\mathfrak{X}_1' \,\&\, \vdash \mathfrak{N}\mathfrak{X}_2'. \tag{3}$$

Therefore

$$\begin{aligned} T\mathfrak{X}' &= T\mathfrak{X}_1'(T\mathfrak{X}_2') && [(2), (3), 9.5], \\ &= \mathfrak{X}_1^*\mathfrak{X}_2^* && [\text{hypothesis of induction}], \\ &= \mathfrak{X}^* && [2.6], \end{aligned}$$

q.e.d.

*13.2. LEMMA. *If* (a) $\vdash \mathfrak{M}$, *and* (b) $\mathfrak{Y}_0 = Z_0$; *then*

$$\vdash T(\mathfrak{R}Y_1 \cdots Y_n).$$

Proof. By (a) and 5.43 it follows that

$$\vdash \mathcal{N}Y_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}Y_n; \tag{1}$$

hence, by 10.22 (2), 7.51, we can conclude (for $j = 1, 2, \cdots, n$) that

$$|\,\mathfrak{Y}_0\,| \cdot \Delta_2 Y_j(\Phi_j\mathfrak{Y}_1 \cdots \mathfrak{Y}_m) = Z_0.$$

Therefore [(b), 7.63]

$$\Delta_2 Y_j(\Phi_j\mathfrak{Y}_1 \cdots \mathfrak{Y}_m) = Z_0.$$

Hence [7.62]

$$Y_j = \Phi_j\mathfrak{Y}_1 \cdots \mathfrak{Y}_m. \tag{2}$$

Thus

$$\begin{aligned} TY_j &= T(\Phi_j\mathfrak{Y}_1 \cdots \mathfrak{Y}_m) && [3.2], \\ &= \begin{bmatrix} T\mathfrak{Y}_1, \cdots, T\mathfrak{Y}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}_j && [13.1, 7.22](^{60}). \end{aligned} \tag{3}$$

But [(a), 5.43] $\vdash TY_j$. Therefore [(3), 3.2]

$$\vdash \begin{bmatrix} T\mathfrak{Y}_1, \cdots, T\mathfrak{Y}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}_j, \tag{4}$$

for $j = 1, 2, \cdots, n$. But by the hypothesis of this section the rule

$$\vdash \mathfrak{X}_1 \,\&\, \cdots \,\&\, \vdash \mathfrak{X}_n \rightarrow_X \vdash \mathfrak{X}_{n+1}$$

is a valid rule of procedure; hence by application of this rule to (4), it follows that

$$\vdash \begin{bmatrix} T\mathfrak{Y}_1, \cdots, T\mathfrak{Y}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}_{n+1}. \tag{5}$$

(60) The hypothesis of 13.1, which in this case is

$$\vdash \mathcal{N}\mathfrak{Y}_1 \,\&\, \cdots \,\&\, \vdash \mathcal{N}\mathfrak{Y}_m,$$

is true by 7.22.

On the other hand, by the definition of ρ (in 10.23) and 7.52,

$$\mathfrak{R}Y_1 \cdots Y_n = |\mathfrak{Y}_0| \cdot (\Phi_{n+1}\mathfrak{Y}_1 \cdots \mathfrak{Y}_m) + \|\mathfrak{Y}_0\| \cdot \mathfrak{Z}_a,$$

where $a \equiv \mathfrak{n}(A_0)$. However, by (b) it follows [6.941] that

$$|\mathfrak{Y}_0| = \mathfrak{Z}_1 \;\&\; \|\mathfrak{Y}_0\| = \mathfrak{Z}_0;$$

and hence [6.91, 6.92] that

$$\mathfrak{R}Y_1 \cdots Y_n = \Phi_{n+1}\mathfrak{Y}_1 \cdots \mathfrak{Y}_m. \tag{6}$$

Therefore

$$T(\mathfrak{R}Y_1 \cdots Y_n) = T(\Phi_{n+1}\mathfrak{Y}_1 \cdots \mathfrak{Y}_m) = \begin{bmatrix} T\mathfrak{Y}_1, \cdots, T\mathfrak{Y}_m \\ X_1, \cdots, X_m \end{bmatrix} \mathfrak{X}_{n+1} \quad [13.1].$$

From this and (5) the lemma follows by 3.2.

*13.3. LEMMA. *If* (a) $\vdash \mathfrak{M}$, *and* (b) $|\mathfrak{Y}_0| = \mathfrak{Z}_0$, *then*

$$\vdash T(\mathfrak{R}Y_1 \cdots Y_n).$$

Proof. From the definition of $\mathfrak{R}$ and (b) it follows by 6.94, 6.92, 6.91 (cf. derivation of 13.2 (6)) that

$$\mathfrak{R}Y_1 \cdots Y_n = \mathfrak{Z}_a,$$

where $a \equiv \mathfrak{n}(A_0)$. Therefore [9.6]

$$T(\mathfrak{R}Y_1 \cdots Y_n) = A_0.$$

The lemma follows from

$$\vdash A_0,$$

which holds by definition of A_0.

13.4. THEOREM. *If* $\mathfrak{R} \equiv \mathfrak{F}(\rho)$ *where* ρ *is defined as in* 10.2 *from the valid rule of procedure*

$$\vdash \mathfrak{X}_1 \;\&\; \cdots \;\&\; \vdash \mathfrak{X}_n \rightarrow_X \vdash \mathfrak{X}_{n+1},$$

then

$$\vdash \mathfrak{N}Y_1 \;\&\; \cdots \;\&\; \vdash \mathfrak{N}Y_n \;\&\; \vdash TY_1 \;\&\; \cdots \;\&\; \vdash TY_n \rightarrow_Y \vdash T(\mathfrak{R}Y_1 \cdots Y_n).$$

Proof. See the introduction to the section.

14. **Formalization of θ; the function Θ.** We turn now to the study of the constant Θ defined by

$$\Theta \equiv \mathfrak{F}(\theta).$$

The principal theorem to be established is that

$$\vdash \mathcal{N}X \rightarrow_X \vdash T(\Theta X),$$

which is proved in 14.6.

Since Θ satisfies a course-of-values recursion, it is natural to approach our theorems by way of a "course-of-values function" $\mathfrak{M}$, such that $\vdash \mathfrak{M}X$ means intuitively "for all $Y \leqq X$, $\vdash T(\Theta X)$." Such an $\mathfrak{M}$ can be formally defined by means of an auxiliary term $\mathfrak{L}$, thus:

$$\mathfrak{L} \equiv \lambda^2 XY.\Lambda(\mathcal{N}Y)(Q(Y \dot{-} X)\mathsf{Z}_0),$$
$$\mathfrak{M} \equiv \lambda X.F(\mathfrak{L}X)T\Theta.$$

Other functions used in this section are symbolized thus:

$$\mathfrak{P}_j \equiv \mathfrak{F}(\pi_j) \qquad (j = 0, 1, \cdots, n),$$
$$\mathfrak{R}_k \equiv \mathfrak{F}(\rho_k) \qquad (k = 0, 1, \cdots, r-1),$$

where ρ_k is the recursive function associated with the kth rule as in 10.3, and

$$\Gamma \equiv \mathfrak{F}(\gamma).$$

The development of the proof is as follows. The idea is to prove by induction that

$$\vdash \mathcal{N}X \rightarrow_X \vdash \mathfrak{M}X.$$

The two parts of this induction are in 14.1 and 14.5. The second part splits into two cases of which the first splits again into two subcases; these cases are taken care of in 14.2–14.3.

An additional, but obvious, property of Θ is proved in 14.7.

14.1. LEMMA. $\vdash \mathfrak{M}\mathsf{Z}_0$.

Proof. Suppose that

$$\vdash \mathfrak{L}\mathsf{Z}_0 Y. \tag{1}$$

Then [5.42] $\vdash \mathcal{N}Y$, and $Y \dot{-} \mathsf{Z}_0 = \mathsf{Z}_0$. Therefore [6.93]

$$Y = \mathsf{Z}_0. \tag{2}$$

On the other hand, by 10.3 (8) and 7.54

$$\Theta\mathsf{Z}_0 = \mathsf{Z}_a,$$

where $a \equiv \mathfrak{n}(A_0)$. Therefore [9.6] $T(\Theta\mathsf{Z}_0) = A_0$. Hence [definition of A_0, 3.2] q.e.d.

$$\vdash T(\Theta\mathsf{Z}_0),$$

and therefore [(2)]

(3) $$\vdash T(\Theta Y).$$

Since (3) has been deduced from the assumption (1), it follows by 3.65 that

$$\vdash F(\mathfrak{L}\mathsf{Z}_0)T\Theta$$
$$= \vdash \mathfrak{M}\mathsf{Z}_0 \qquad \text{[definition]},$$

q.e.d.

*14.2. LEMMA. *If* (a) $\vdash \mathcal{N}X$, (b) $\vdash \mathfrak{M}X$, (c) $\mathfrak{P}_0(\mathfrak{S}X) = \mathsf{Z}_h$ $(0 \leqq h < r)$; *then*

$$\vdash T(\Theta(\mathfrak{S}X)).$$

Proof. Let $\mathcal{Y} \equiv \mathfrak{S}X$. Then, by 10.3 (9) and 7.51 we have

(1) $$\Theta\mathcal{Y} = \sum_{k=0}^{r-1} \left| \Delta_2(\mathfrak{P}_0\mathcal{Y})\mathsf{Z}_k \right| \cdot \mathfrak{R}_k(\Theta(\mathfrak{P}_1\mathcal{Y})) \cdots (\Theta(\mathfrak{P}_n\mathcal{Y}))$$
$$+ \left\| \mathfrak{P}_0\mathcal{Y} \dot{-} \mathsf{Z}_{r-1} \right\| \cdot \mathcal{A}(\Gamma\mathcal{Y}).$$

To this now apply (c), using the fact that [6.95]

$$\left| \Delta_2\mathsf{Z}_h\mathsf{Z}_k \right| = \mathsf{Z}_0 \qquad \text{for } h \neq k,$$
$$\left| \Delta_2\mathsf{Z}_h\mathsf{Z}_h \right| = \mathsf{Z}_1,$$

and [6.93]

$$\left\| \mathsf{Z}_h \dot{-} \mathsf{Z}_{r-1} \right\| = \mathsf{Z}_0.$$

Then [6.91, 6.92]

(2) $$\Theta\mathcal{Y} = \mathfrak{R}_h(\Theta(\mathfrak{P}_1\mathcal{Y})) \cdots (\Theta(\mathfrak{P}_n\mathcal{Y})).$$

Now by 10.3 (10) and 7.51 it follows that

$$\mathfrak{P}_j\mathcal{Y} \dot{-} X = \mathsf{Z}_0 \qquad (j = 1, 2, \cdots, n),$$

also [7.22, (a), 3.62, 3.63] $\vdash \mathcal{N}(\mathfrak{P}_j\mathcal{Y})$. Therefore [5.42, definition of $\mathfrak{L}$]

$$\vdash \mathfrak{L}X(\mathfrak{P}_j\mathcal{Y}).$$

Hence [(b), 3.63]

(3) $$\vdash T(\Theta(\mathfrak{P}_j\mathcal{Y})).$$

But [7.22, (a), 3.62, 3.63]

(4) $$\vdash \mathcal{N}(\Theta(\mathfrak{P}_j\mathcal{Y})).$$

Therefore [(2), (3), (4), 13.4]

$$\vdash T(\Theta\mathcal{Y}),$$

q.e.d.

*14.3. LEMMA. *If* (a) $\vdash \mathcal{N}X$, (b) $\vdash \mathfrak{M}X$, (c) $\left| \mathfrak{P}_0(\mathfrak{S}X) \dot{-} \mathsf{Z}_{r-1} \right| = \mathsf{Z}_0$, *then* $\vdash T(\Theta(\mathfrak{S}X))$.

Proof. As in 14.2 we conclude from 10.3 (10) and (c), using theorems of 6.9, 7.5 and 7.6, that

$$\Theta(\mathfrak{S}\,X) = \mathscr{A}(\Gamma(\mathfrak{S}\,X)).$$

By (a), 3.62, 3.63, 7.22 it follows that

$$\vdash \mathcal{N}(\Gamma(\mathfrak{S}\,X)).$$

Therefore [§12]

$$T(\mathscr{A}(\Gamma(\mathfrak{S}\,X))).$$

The lemma follows by 3.2.

*14.4. LEMMA. *If* (a) $\vdash \mathcal{N}X$, (b) $\vdash \mathfrak{M}X$; *then*

$$\vdash T(\Theta(\mathfrak{S}\,X)).$$

Proof. [14.2, 14.3, 7.22, 11.3].

*14.5. LEMMA. *If* (a) $\vdash \mathcal{N}X$, *and* (b) $\vdash \mathfrak{M}X$, *then* $\vdash \mathfrak{M}(\mathfrak{S}\,X)$.

Proof. Suppose that

(1) $$\vdash \mathfrak{L}(\mathfrak{S}\,X)Y.$$

Then [5.42]

(2) $$\vdash \mathcal{N}Y,$$

(3) $$Y \dot{-} \mathfrak{S}\,X = \mathfrak{Z}_0.$$

Now in 11.2 it has been shown that from (a) and (2) it follows that

$$|\,Y \dot{-} \mathfrak{S}\,X| \cdot \|Y \dot{-} X\| \cdot \Delta_2 Y(\mathfrak{S}\,X) = \mathfrak{Z}_0.$$

Therefore [(3), 7.63]

(4) $$\|Y \dot{-} X\| \cdot \Delta_2 Y(\mathfrak{S}\,X) = \mathfrak{Z}_0.$$

Suppose then that

(5) $$Y \dot{-} X = \mathfrak{Z}_0.$$

Then [(2), 5.42] $\vdash \mathfrak{L}XY$. Therefore [(b), 3.63]

(6) $$\vdash T(\Theta Y),$$

on the assumption that (1) and (5) hold. On the other hand, suppose that

(7) $$|\,Y \dot{-} X| = \mathfrak{Z}_0.$$

Then [(4), 7.63, 7.62] $Y = \mathfrak{S}\,X$. Therefore [14.4]

$$\vdash T(\Theta Y).$$

This has been shown to follow from (1) and (5) and also from (1) and (7). Hence, by 11.1, it follows from (1) only. The lemma now follows by definition of $\mathfrak{M}$ and 3.65.

14.6. THEOREM. $\vdash \mathcal{N}X \rightarrow_X \vdash T(\Theta X)$.

Proof. From 14.1, 14.5 it follows by 5.44 that

$$\vdash \mathcal{N}X \rightarrow_X \vdash \mathfrak{M}X.$$

Therefore $[3.63, 5.42] \vdash \mathcal{N}X \;\&\; \vdash \mathcal{N}Y \;\&\; Y \doteq X = \mathsf{Z}_0 \rightarrow_{XY} \vdash T(\Theta Y)$. In this substitute X for Y and we have the theorem.

14.7. THEOREM. *If* $\vdash \mathfrak{A}$, *then there exists an* n *such that* $\mathfrak{A} = T(\Theta \mathsf{Z}_n)$.

Proof. Let $a \equiv \mathfrak{n}(\mathfrak{A})$. From 10.3 it is evident that if a proof of $\vdash \mathfrak{A}$ is at hand we can find an n such that $a = \theta(n)$. Then $[7.54]$ $\mathsf{Z}_a = \Theta \mathsf{Z}_n$. Therefore $[9.6]$

$$\mathfrak{A} = T\mathsf{Z}_a = T(\Theta \mathsf{Z}_n).$$

PART V. FINAL DEDUCTION

15. **Proof of inconsistency.** We are now in a position to set up the Richard paradox within the system. The following explanations may help in the understanding of the derivation. The term $\mathfrak{H}$, which in the first two lemmas is taken as unspecified, is defined to be $F\mathcal{N}\mathcal{N}$ in 15.3; so that $\mathfrak{H}$ represents the property of being a numerical function. Then the function $\omega(x)$ introduced in 15.1 is a recursive enumeration (with repetitions, of course) of the numbers associated with all the numerical functions. On the other hand $\mathfrak{U}$, introduced in 15.2, is an enumeration within the system of the numerical functions themselves; thus $\mathfrak{U}$ represents the sequence $\phi_1(x), \phi_2(x), \cdots$ of §1. The diagonal function, $f(x)$ in §1, is the $\mathfrak{G}$ of §15.3; and the number of $\mathfrak{G}$ in the enumeration $\mathfrak{U}$ is Z_g.

15.1. LEMMA. *If* $\mathfrak{H}$, $\mathfrak{A}$ *are terms such that*

$$\vdash \mathfrak{H}\mathfrak{A}, \tag{1}$$

then there exist recursive functions $\omega(x)$, $\sigma'(x)$ *such that*:

(i) *if* $\mathfrak{X}$ *is any term such that we have a proof that*

$$\vdash \mathfrak{H}\mathfrak{X}, \tag{2}$$

then we can find an integer n *such that*

$$\mathfrak{n}(\mathfrak{X}) = \omega(n);$$

(ii) *we have further*

$$\mu(\mathfrak{n}(\mathfrak{H}), \omega(x)) = \theta(\sigma'(x)). \tag{3}$$

Remark. Property (ii) is a kind of converse of (i); for if $\mathfrak{n}(\mathfrak{X})=\omega(n)$,then by (ii) and 3.521, $\mathfrak{n}(\mathfrak{H}\mathfrak{X})=\theta(\sigma'(n))$, and so, by the construction of θ [10.3], $\vdash \mathfrak{H}\mathfrak{X}$. Thus ω enumerates all and only those integers which are numbers of terms $\mathfrak{X}$ for which (2) holds.

Proof. Let $h \equiv \mathfrak{n}(\mathfrak{H})$,

$$a \equiv \mathfrak{n}(\mathfrak{A}), \tag{4}$$

and let b be the number, obtained from a proof of (1) as in 10.3, such that [3.521]

$$\mathfrak{n}(\mathfrak{H}\mathfrak{A}) = \mu(h, a) = \theta(b). \tag{5}$$

Let $\sigma'(x)$ be the recursive function such that(81)

$$\sigma'(x) = \begin{cases} x & \text{if } \mu_1(\theta(x)) = h, \\ b & \text{if } \mu_1(\theta(x)) \neq h; \end{cases} \tag{6}$$

and then let

$$\omega(x) \equiv \mu_2(\theta(\sigma'(x))). \tag{7}$$

Now suppose that (2) holds and let

$$m \equiv \mathfrak{n}(\mathfrak{X}).$$

Let n be such that

$$\mathfrak{n}(\mathfrak{H}\mathfrak{X}) = \mu(h, m) = \theta(n). \tag{8}$$

Then [8.1 (3)] $\mu_1(\theta(n))=h$. Therefore [(6)]

$$\sigma'(n) = n. \tag{9}$$

Therefore

$$\begin{aligned} \omega(n) &= \mu_2(\theta(n)) && [(7)], \\ &= m && [(8), 8.1, (4)], \end{aligned}$$

which proves (i).

To prove (ii), we first show that

$$\mu_1(\theta(\sigma'(x))) = h. \tag{10}$$

If $\mu_1(\theta(x))=h$, this follows at once by (6); if not, then it follows from (6) and (5) by 8.1 (3). By 6.86 this establishes (10) as a theorem in r. But since [(4), 3.5] $h>0$, it then follows that [8.1, (8)]

(81) An explicit definition of σ' is

$$\sigma'(x) = \mathfrak{B}|\mathfrak{a}| \cdot x + \|\mathfrak{a}\| \cdot b$$

where $\mathfrak{a} = \delta_2(h, \mu_1(\theta(x)))$.

$$\mu_3(\theta(\sigma'(x))) = 0.$$

Therefore [8.1, (7)]

$$\theta(\sigma'(x)) = \mu(h, \omega(x)),$$

q.e.d.

15.2. LEMMA. *If* $\mathfrak{H}$, $\mathfrak{A}$, σ', ω *are as in* 15.1, *and if*

(1) $$\Omega \equiv \mathfrak{F}(\omega),$$

(2) $$\mathfrak{U} \equiv \lambda X.T(\Omega X);$$

then

(i) *if* $\mathfrak{X}$ *is any term such that*

(3) $$\vdash \mathfrak{H}\mathfrak{X},$$

then from a proof of (3) *we can find an* n *such that*

$$\mathfrak{X} = \mathfrak{U}\mathfrak{Z}_n;$$

(ii) $\vdash F\mathfrak{N}\mathfrak{H}\mathfrak{U}$.

Remark. Here again (ii) implies the converse of (i) [3.63, 5.27]. $\mathfrak{U}$ gives a formal enumeration, not of numbers of the terms satisfying (3), but of the terms themselves.

Proof. Suppose that (3) holds and set

$$m \equiv \mathfrak{n}(\mathfrak{X}).$$

Then by 15.1 we can find an n such that

$$m = \omega(n).$$

Therefore [7.53]

$$\mathfrak{Z}_m = \Omega\mathfrak{Z}_n.$$

Hence

$$\mathfrak{X} = T\mathfrak{Z}_m = T(\Omega\mathfrak{Z}_n) \quad [9.6],$$
$$= \mathfrak{U}\mathfrak{Z}_n \quad [(2)].$$

This proves (i).

Let us turn to (ii). As in 15.1 let

$$h \equiv \mathfrak{n}(\mathfrak{H}), \qquad \Sigma \equiv \mathfrak{F}(\sigma').$$

Then by 15.1 (ii) and 7.51

$$\vdash \mathfrak{N}X \to_X \Theta(\Sigma X) = M\mathfrak{Z}_h(\Omega X).$$

Therefore

$$
\begin{aligned}
(4)\qquad \vdash \mathcal{N}X \rightarrow_X T(\Theta(\Sigma X)) &= T(M\mathfrak{Z}_h(\Omega X)) \\
&= T\mathfrak{Z}_h(T(\Omega X)) && [9.5, 5.27, 7.22], \\
&= \mathfrak{H}(\mathfrak{U}X) && [9.6, (2)].
\end{aligned}
$$

But

$$
\begin{aligned}
(5)\qquad \vdash \mathcal{N}X &\rightarrow_X \vdash \mathcal{N}(\Sigma X) && [7.22], \\
&\rightarrow_X \vdash T(\Theta(\Sigma X)) && [14.6].
\end{aligned}
$$

Therefore [(4), (5)]

$$\vdash \mathcal{N}X \rightarrow_X \vdash \mathfrak{H}(\mathfrak{U}X).$$

(ii) now follows by 3.65.

15.3. LEMMA. *Let* $\mathfrak{U}$ *be defined as in* 15.2 *with*

$$\mathfrak{H} \equiv F\mathcal{N}\mathcal{N}, \qquad \mathfrak{A} \equiv \mathfrak{S},$$

and let $\mathfrak{G} \equiv \lambda X.\mathfrak{S}(\mathfrak{U}XX)$; *then we may find an integer* g *such that*

(i) $\mathfrak{G} = \mathfrak{U}\mathfrak{Z}_g$, *and*

(ii) $\mathfrak{U}\mathfrak{Z}_g\mathfrak{Z}_g = \mathfrak{S}(\mathfrak{U}\mathfrak{Z}_g\mathfrak{Z}_g)$.

Proof. The only hypothesis on $\mathfrak{H}$ and $\mathfrak{A}$ in 15.1 and 15.2 was 15.1 (1); and this is fulfilled for the present specialized $\mathfrak{H}$ and $\mathfrak{A}$ by 3.62. Hence, by 15.2 (ii),

$$\vdash F\mathcal{N}(F\mathcal{N}\mathcal{N})\mathfrak{U}.$$

Therefore

$$
\begin{aligned}
(1)\qquad \vdash \mathcal{N}X &\rightarrow_X \vdash F\mathcal{N}\mathcal{N}(\mathfrak{U}X) && [3.63], \\
&\rightarrow_X \vdash \mathcal{N}(\mathfrak{U}XX) && [3.63], \\
&\rightarrow_X \vdash \mathcal{N}(\mathfrak{S}(\mathfrak{U}XX)) && [3.62, 3.63], \\
&\rightarrow_X \vdash \mathcal{N}(\mathfrak{G}X) && [\text{definition of } \mathfrak{G}].
\end{aligned}
$$

Therefore [3.65]

$$\vdash F\mathcal{N}\mathcal{N}\mathfrak{G}.$$

From this and 15.2 (i) it follows that there exists an integer g such that (i) holds. From (i) we obtain (ii) by applying both sides of (i) to $\mathfrak{Z}_g$ and using the definition of $\mathfrak{G}$ in connection with 3.22.

15.4. THEOREM. $\mathfrak{Z}_0 = \mathfrak{Z}_1$.

Proof. Let $\mathfrak{U}$, $\mathfrak{G}$ and g be as in 15.3 and let

$$\mathfrak{B} \equiv \mathfrak{U}\mathfrak{Z}_g\mathfrak{Z}_g.$$

Then [15.3 (ii)] $\mathfrak{B} = \mathfrak{S}\mathfrak{B}$. Therefore

(1) $$\mathfrak{B} \dot{-} \mathfrak{B} = \mathfrak{S}\mathfrak{B} \dot{-} \mathfrak{B}.$$

But [15.3 (1)] $\vdash \mathfrak{N}\mathfrak{B}$. Therefore [6.93]

$$\mathfrak{B} \dot{-} \mathfrak{B} = \mathfrak{Z}_0,$$

$$\mathfrak{S}\mathfrak{B} \dot{-} \mathfrak{B} = \mathfrak{Z}_1.$$

The theorem now follows by (1).

15.5. THEOREM. *If* $\mathfrak{A}$ *is a term, then* $\vdash \mathfrak{A}$.

Proof. By 15.4 $\mathfrak{Z}_0 = \mathfrak{Z}_1$. Therefore [3.2]

$$D_2 A_0 \mathfrak{A} \mathfrak{Z}_0 = D_2 A_0 \mathfrak{A} \mathfrak{Z}_1.$$

Therefore [5.31, 5.32]

$$A_0 = \mathfrak{A}.$$

But [3.14] $\vdash A_0$, hence [3.2]

$$\vdash \mathfrak{A},$$

q.e.d.

PENNSYLVANIA STATE COLLEGE,
STATE COLLEGE, PA.

A REFINEMENT OF AN INEQUALITY OF THE BROTHERS MARKOFF

BY

R. J. DUFFIN AND A. C. SCHAEFFER

INTRODUCTION

Suppose that $f(x)$ is a polynomial of degree n such that $|f(x)| \leqq 1$ in the interval $(-1, 1)$. Then according to a theorem(1) of A. Markoff $|f'(x)| \leqq n^2$ in the same interval. This bound is actually attained by the Tchebychef polynomial $T_n(x) = \cos(n \cos^{-1} x)$. A later theorem(2) of W. Markoff gives the best possible bounds for the higher derivatives of $f(x)$

$$|f^{(p)}(x)| \leqq \frac{n^2(n^2 - 1^2) \cdots (n^2 - (p-1)^2)}{1 \cdot 3 \cdot 5 \cdots (2p-1)}, \qquad -1 \leqq x \leqq 1.$$

We show in this paper that the condition $|f(x)| \leqq 1$ in the whole interval is unnecessarily restrictive. Instead it is only necessary to assume that $|f(x)| \leqq 1$ at the $n+1$ points $x = \cos(k\pi/n)$; $k = 0, 1, 2, \cdots, n$, and the conclusions of Markoffs' theorem are unchanged. These are exactly the points at which $T_n(x)$ attains its maximum in the interval $(-1, 1)$.

In a previous paper(3) we showed that under the assumptions of Markoffs' theorem all derivatives of $f(x)$ must satisfy

$$\left|\frac{d^p}{dx^p} f(x)\right| \leqq \left|\frac{d^p}{dx^p} e^{in \arccos x}\right|, \qquad -1 \leqq x \leqq 1. \tag{1}$$

For $p = 1$ this reduces to an inequality of S. Bernstein

$$|f'(x)| \leqq n(1 - x^2)^{-1/2}.$$

Using (1) it is possible to deduce W. Markoff's result. However the proof which is given here is simpler. Actually (1) is not true under the weaker conditions which are assumed in the present paper. The W. Markoff inequalities are obtained here as a consequence of a more general inequality concerning Lagrange interpolation.

Presented to the Society, September 8, 1939, received by the editors September 10, 1940.

(1) A. Markoff, *Sur une question posée par Mendeleieff*, Bulletin of the Academy of Sciences of St. Petersburg, vol. 62 (1889), pp. 1–24.

(2) W. Markoff, *Über Polynome, die in einen gegebenen Intervalle möglichst wenig von Null abweichen*, Mathematische Annalen, vol. 77 (1916), pp. 213–258. The original appeared in Russian in 1892.

(3) R. J. Duffin and A. C. Schaeffer, *On some inequalities of S. Bernstein and W. Markoff for derivatives of polynomials*, Bulletin of The American Mathematical Society, vol. 44 (1938), pp. 289–297.

I. An inequality of interpolation

Suppose $g(z)$ is a polynomial of degree n with n distinct real zeros lying in the interval (a, b) of the real axis,

$$g(z) = c\prod_{\nu=1}^{n}(z-\lambda_\nu), \qquad c \neq 0;$$

$$g'(z) = \sum_{\nu=1}^{n}\frac{g(z)}{z-\lambda_\nu}.$$

We seek some way of determining the maximum of the non-analytic function

$$M(x) = \sum_{\nu=1}^{n}\left|\frac{g(x)}{x-\lambda_\nu}\right|$$

when x lies in the interval (a, b). By making a certain hypothesis concerning the behavior of $g(z)$ in the strip $a \leqq \Re(z) \leqq b$ (inequality (4) of Theorem I) we obtain the elegant solution

$$M(x) \leqq M(b) = |g'(b)|, \qquad a \leqq x \leqq b.$$

This result is equivalent to the following statement: suppose that $f(z)$ is any polynomial of degree n whose derivative satisfies the condition $|f'(x)| \leqq |g'(x)|$ at the zeros of $g(x)$; then (with the same restrictions as before regarding $g(z)$)

$$|f'(x)| \leqq |g'(b)|, \qquad a \leqq x \leqq b.$$

The equivalence is an immediate consequence of the Lagrange interpolation formula; we have

$$f'(z) = \sum_{\nu=1}^{n}\frac{f'(\lambda_\nu)}{g'(\lambda_\nu)}\,\frac{g(z)}{z-\lambda_\nu}.$$

Hence, under the conditions of the last statement, if x is any point in (a, b)

$$\max|f'(x)| = \sum_{\nu=1}^{n}\left|\frac{g(x)}{x-\lambda_\nu}\right|.$$

We shall also be concerned with finding bounds for the higher derivatives of the polynomial $f(z)$, and to accomplish this we need the following lemma.

Lemma I. *If $g(z)=c\prod_{\nu=1}^{n}(z-\lambda_\nu)$, $c\neq 0$, is a polynomial of degree n with n distinct real zeros and if $f(z)$ is a polynomial of degree n such that*

$$|f'(\lambda_\nu)| \leqq |g'(\lambda_\nu)|; \qquad \nu = 1, 2, 3, \cdots, n,$$

then for $p=1, 2, 3, \cdots, n$ we must have at the zeros of $g^{(p-1)}(z)$

$$|f^{(p)}(z)| \leqq |g^{(p)}(z)|. \tag{2}$$

Proof. For $p=1$ inequality (2) is simply a restatement of the conditions of the lemma; so consider the case $p=2$. The Lagrange interpolation formula gives

$$\frac{f'(z)}{g(z)} = \sum_{\nu=1}^{n} \frac{f'(\lambda_\nu)}{g'(\lambda_\nu)} \frac{1}{z-\lambda_\nu} = \sum_{\nu=1}^{n} \frac{\delta_\nu}{z-\lambda_\nu} \tag{3}$$

where by the hypotheses of the lemma, $|\delta_\nu| \leqq 1$. There is a similar expression for $g'(z)$ in which each δ_ν is equal to 1. Differentiating (3) we obtain

$$\{f''(z)g(z) - f'(z)g'(z)\}/\{g(z)\}^2 = -\sum_{\nu=1}^{n} \frac{\delta_\nu}{(z-\lambda_\nu)^2}.$$

Thus at the points where $g'(z)=0$ we have

$$\left|\frac{f''(z)}{g(z)}\right| = \left|\sum_{\nu=1}^{n} \frac{\delta_\nu}{(z-\lambda_\nu)^2}\right| \leqq \sum_{\nu=1}^{n} \frac{1}{(z-\lambda_\nu)^2} = \left|\frac{g''(z)}{g(z)}\right|$$

and it follows that Lemma I is true for $p=2$.

The proof for $p>2$ is by induction. Let $|f^{(p)}(x)| \leqq |g^{(p)}(x)|$ at the zeros of $g^{p-1}(x)$ (they are real and distinct). By applying the previous argument to $f^{(p)}(x)$ and $g^{(p)}(x)$ instead of $f'(x)$ and $g'(x)$ we obtain

$$|f^{(p+1)}(x)| \leqq |g^{(p+1)}(x)|$$

at the zeros of $g^{(p)}(x)$. This completes the induction.

THEOREM I. *Let $g(z)$ be a polynomial of degree n with n distinct real zeros to the left of the point b on the real axis, and suppose that in a strip of the complex plane it satisfies the inequality*

$$|g(x+iy)| \leqq |g(b+iy)|; \qquad a \leqq x \leqq b, \ -\infty < y < \infty. \tag{4}$$

Then if $f(z)$ is a polynomial of degree n with real coefficients such that

$$|f'(x)| \leqq |g'(x)| \quad \textit{wherever} \quad g(x) = 0 \tag{5}$$

the derivatives of $f(z)$ and $g(z)$ must satisfy

$$|f^{(p)}(x+iy)| \leqq |g^{(p)}(b+iy)|; \qquad a \leqq x \leqq b, \ -\infty < y < \infty, \ p = 1, 2, 3, \cdots, n. \tag{6}$$

Proof. To prove Theorem I we shall first show that at a fixed point x_1+iy_1 in the strip

$$|f'(x_1+iy_1)| \leqq |g'(b+iy_1)|. \tag{7}$$

Let $g(z)=c\prod_{\nu=1}^{n}(z-\lambda_\nu)$, $c\neq 0$, where without loss of generality we suppose that c is real.

Let $h(z)$ be a new polynomial with the same leading coefficient as $g(z)$, and whose roots are obtained by reflecting about x_1 those roots of $g(z)$ which lie to the right of x_1. Thus

$$h(z) = c\prod_{\nu=1}^{n}(z-\beta_\nu)$$

where

$$\beta_\nu = \begin{cases} 2x_1 - \lambda_\nu, & \text{if } \lambda_\nu > x_1, \\ \lambda_\nu, & \text{if } \lambda_\nu \leqq x_1. \end{cases}$$

Then on the line $z=x_1+iy$, $-\infty<y<\infty$, we must have $|z-\beta_\nu| = |z-\lambda_\nu|$ so

$$|h(x_1+iy)| = |g(x_1+iy)|. \tag{8}$$

We now show that at the point x_1+iy_1 the derivative of $h(z)$ is at least as large as the derivative of $f(z)$. Whether or not the roots of $h(z)$ are distinct we may write

$$\frac{h'(z)}{h(z)} = \sum_{\nu=1}^{n}\frac{1}{z-\beta_\nu} \tag{9}$$

and $f'(z)$ is given by the interpolation formula (3), where now the numbers δ_ν satisfy $-1\leqq\delta_\nu\leqq 1$. Comparing the right-hand sides of (3) and (9) we have

$$\begin{aligned}\left|\sum_{\nu=1}^{n}\frac{\delta_\nu}{x_1+iy_1-\lambda_\nu}\right| &= \left|\sum\frac{\delta_\nu(x_1-\lambda_\nu)}{(x_1-\lambda_\nu)^2+y_1^2} - i\sum\frac{y_1\delta_\nu}{(x_1-\lambda_\nu)^2+y_1^2}\right| \\ &\leqq \left|\sum\frac{x_1-\beta_\nu}{(x_1-\beta_\nu)^2+y_1^2} - i\sum\frac{y_1}{(x_1-\beta_\nu)^2+y_1^2}\right| \\ &= \left|\sum\frac{1}{x_1-\beta_\nu+iy_1}\right|\end{aligned}$$

since by construction $|x_1-\lambda_\nu| = x_1-\beta_\nu$. Since $g(z)$ and $h(z)$ have the same absolute magnitude at x_1+iy_1 it follows that $|h'(x_1+iy_1)| \geqq |f'(x_1+iy_1)|$.

Let α, $|\alpha|<1$, be any complex constant and let

$$\phi(z) = g(z) - \alpha h(z+x_1-b).$$

Let Γ be the simple closed curve consisting of a segment of the line $z=b+iy$ and the portion of a circle with center at b and radius ρ which lies to the right of this line. Relations (4) and (8) show that on the straight line segment of Γ

$$|g(z)| > |\alpha h(z+x_1-b)|.$$

If ρ is sufficiently large the same inequality is true for the circular portion of Γ since $g(z)$ and $h(z)$ have the same leading coefficient. Thus Rouché's theorem states that $\phi(z)$ and $g(z)$ have the same number of zeros inside Γ. We conclude that $\phi(z)$ has no zeros on or to the right of the line $b+iy$. The

last statement, according to a theorem of Gauss, implies that $\phi'(z)$ has no zeros on the line $b+iy$. Thus for $|\alpha|<1$

$$g'(b+iy_1)-\alpha h'(x_1+iy_1)\neq 0.$$

It follows immediately that

$$|g'(b+iy_1)|\geqq|h'(x_1+iy_1)|\geqq|f'(x_1+iy_1)|.$$

This proves the theorem for $p=1$. We turn now to the case $p>1$.

Applying Theorem I for $p=1$ when $f(z)$ is the function $g(z)$ itself inequality (6) shows that

$$|g'(x+iy)|\leqq|g'(b+iy)|;\qquad a\leqq x\leqq b,\ -\infty<y<\infty.$$

Thus $g'(z)$ satisfies all the requirements which in Theorem I are imposed on the interpolating polynomial $g(z)$ (with n replaced by $n-1$), and by Lemma I, $|f''(x)|\leqq|g''(x)|$ at the $n-1$ zeros of $g'(x)$. Applying Theorem I to $f'(x)$ instead of $f(x)$ in the case $p=1$, for which it has already been proved true, we have

$$|f''(x+iy)|\leqq|g''(b+iy)|;\qquad a\leqq x\leqq b,\ -\infty<y<\infty,$$

which proves that inequality (6) is true for $p=2$. Repetition of this argument completes the proof for larger values of p.

REMARK TO THEOREM I. *If $f(z)$ and $g(z)$ satisfy the same conditions as in Theorem* I *except the condition that $f(x)$ has real coefficients, then for* $p=1, 2, 3, \cdots, n$

$$|f^{(p)}(x)|\leqq|g^{(p)}(b)|,\qquad a\leqq x\leqq b.$$

Proof. After differentiating (3) $p-1$ times we obtain

$$f^{(p)}(x)=\sum_{\nu=1}^{n}\delta_\nu\frac{d^{p-1}}{dx^{p-1}}\left(\frac{g(x)}{x-\lambda_\nu}\right)$$

where $|\delta_\nu|=|f'(\lambda_\nu)/g'(\lambda_\nu)|\leqq 1$. It is evident that if x is any point in the interval (a, b), $f^{(p)}(x)$ attains its maximum when $\delta_\nu=\pm 1$. The Remark then follows from Theorem I.

II. A PROPERTY OF TCHEBYCHEF POLYNOMIALS

The Tchebychef polynomials $T_n(x)$ which are defined by

$$T_n(x)=\cos n\theta,\qquad x=\cos\theta,$$

are polynomials in x of degree n and have n real distinct roots in the interval $(-1, 1)$. They may be written

$$T_n(z)=c\prod_{\nu=1}^{n}(z-\cos\theta_\nu)\tag{10}$$

where $\theta_\nu=(\nu-\frac{1}{2})\pi/n$; $\nu=1, 2, \cdots, n$ and $c=2^{n-1}$, and they satisfy the differential equation

$$(1-z^2)T_n''(z)-zT_n'(z)+n^2T_n(z)=0. \tag{11}$$

It will be shown that $T_n(z)$ satisfies the conditions which in Theorem I are imposed on the interpolating polynomial $g(z)$ where we take $a=-1$, $b=1$. The proof depends on the following lemma.

LEMMA II. *Let $\{\alpha_i\}$ be a sequence of $2n$ non-negative numbers and let $\{\alpha_i'\}$ be a rearrangement of this sequence according to magnitude,*

$$\alpha_1' \geqq \alpha_2' \geqq \alpha_3' \geqq \cdots \geqq \alpha_{2n}' \geqq 0.$$

Then if $t \geqq 0$

$$(\alpha_1\alpha_2+t)(\alpha_3\alpha_4+t)\cdots(\alpha_{2n-1}\alpha_{2n}+t) \tag{12}$$

is not greater than

$$(\alpha_1'\alpha_2'+t)(\alpha_3'\alpha_4'+t)\cdots(\alpha_{2n-1}'\alpha_{2n}'+t).$$

Proof. If α_1 and α_3 are as large as any of the α's then

$$(\alpha_1\alpha_3+t)(\alpha_2\alpha_4+t)-(\alpha_1\alpha_2+t)(\alpha_3\alpha_4+t)=t(\alpha_1-\alpha_4)(\alpha_3-\alpha_2)\geqq 0.$$

This shows that the α's in (12) can be rearranged so that the two largest occur in the same factor without decreasing (12). Then the two largest of the remaining α's may be brought into the same factor without decreasing (12). And so on.

LEMMA III. *The Tchebychef polynomials satisfy the inequality*

$$|T_n(x+iy)| \leqq |T_n(1+iy)|; \qquad -1 \leqq x \leqq 1, \ -\infty < y < \infty.$$

Proof. First,

$$|T_n(x+iy)|^2 = c^2\prod_{\nu=1}^{n}\{(x-\cos\theta_\nu)^2+y^2\}.$$

If $-1\leqq x\leqq 1$ we may write $x=\cos\theta$, θ real. Then we obtain

$$|T_n(x+iy)|^2 = c^2\prod_{\nu=1}^{n}\{\tfrac{1}{4}|e^{i\theta}-e^{-i\theta_\nu}|^2|e^{i\theta}-e^{i\theta_\nu}|^2+y^2\}. \tag{13}$$

Geometrically, $e^{\pm i\theta_\nu}$, $\nu=1, 2, 3, \cdots, n$, represent $2n$ points equally distributed about the unit circle. Connect these points by chords to the point $e^{i\theta}$. Then the lengths of these $2n$ chords are given by $|e^{i\theta}-e^{\pm i\theta_\nu}|$. If θ is increased or decreased by any multiple of π/n we obtain a new set of chords but the aggregate of their lengths is unchanged.

Choose ϕ such that

$$\phi \equiv \theta (\text{mod } \pi/n), \qquad -\pi/(2n) \leqq \phi \leqq \pi/(2n).$$

If $x^* = \cos \phi$ then

$$T_n(x^* + iy) = c^2 \prod_{\nu=1}^{n} \{\tfrac{1}{4} \left| e^{i\phi} - e^{-i\theta_\nu} \right|^2 \left| e^{i\phi} - e^{i\theta_\nu} \right|^2 + y^2\} \tag{14}$$

where the numbers $\left| e^{i\phi} - e^{\pm i\theta_\nu} \right|^2$ are simply a rearrangement of the numbers $\left| e^{i\theta} - e^{\pm i\theta_\nu} \right|^2$ appearing in (13).

We observe that Lemma II is directly applicable to the products (14) and (13). Since, as seen geometrically, the pairing of the numbers $\left| e^{i\phi} - e^{\pm i\theta_\nu} \right|^2$ in (14) is by magnitude, we have

$$\left| T_n(x + iy) \right|^2 \leqq \left| T_n(x^* + iy) \right|^2.$$

By construction $x^* + iy$ lies in the strip $\cos \theta_1 \leqq x \leqq 1$ where $\cos \theta_1$ is the right most zero of $T_n(x)$. Thus the distance from any roots to $x^* + iy$ is no greater than the distance from the same root to $1 + iy$, so (10) gives

$$\left| T_n(x^* + iy) \right| \leqq \left| T_n(1 + iy) \right|.$$

This completes the proof of Lemma III.

Pólya and Szegö prove(4) that if $P(x)$ is a polynomial of degree $n-1$ such that

$$(1 - x^2)^{1/2} \left| P(x) \right| \leqq 1, \qquad -1 \leqq x \leqq 1, \tag{15}$$

then

$$\left| P(x) \right| \leqq n, \qquad -1 \leqq x \leqq 1. \tag{16}$$

We show that if (15) is satisfied at the n points $x = \cos(\pi(\nu - \frac{1}{2})/n)$; $\nu = 1, 2, \cdots, n$ then (16) is still true. We use Theorem I with $g(x) = T_n(x)$. When $T_n(x) = 0$, $T_n'(x) = n(1-x^2)^{-1/2}$ so we may write $f'(x) = nP(x)$. Thus for $-1 \leqq x \leqq 1$,

$$\left| nP(x) \right| \leqq T_n'(1) = n^2.$$

III. Markoffs' inequalities

Lemma IV. *Let $f(x)$ be a polynomial of degree n satisfying $|f(x)| \leqq 1$ at $x = \cos(\nu\pi/n)$; $\nu = 0, 1, 2, 3, \cdots, n$. Then*

$$\left| f'(x) \right| \leqq n(1 - x^2)^{-1/2} \quad \text{at} \quad x = \cos(\pi(\nu - \tfrac{1}{2})/n); \qquad \nu = 1, 2, 3, \cdots, n.$$

The equality can occur only if $f(x) \equiv \gamma T_n(x)$ for some constant γ, $|\gamma| = 1$.

(4) G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, p. 90, Problem 80.

Proof. Let

$$\psi(x) = (1 - x^2)T_n'(x) = a\prod_{\nu=0}^{n}(x - \cos(\nu\pi/n)).$$

Then by (11) $\psi'(x) = -xT_n'(x) - n^2T_n(x)$. Differentiating the Lagrange interpolation formula for $f(x)$ gives, if $\lambda_\nu = \cos(\nu\pi/n)$,

$$f'(x) = \sum_{\nu=0}^{n} \frac{f(\lambda_\nu)}{\psi'(\lambda_\nu)} \frac{(x - \lambda_\nu)\psi'(x) - \psi(x)}{(x - \lambda_\nu)^2}.$$

At the zeros of $T_n(x)$ this reduces to

$$f'(x) = -T_n'(x)\sum_{0}^{n} \frac{f(\lambda_\nu)}{\psi'(\lambda_\nu)} \frac{1 - x\lambda_\nu}{(x - \lambda_\nu)^2} \tag{17}$$

since at these points

$$\begin{aligned}(x - \lambda_\nu)\psi'(x) - \psi(x) &= -x(x - \lambda_\nu)T_n'(x) - (1 - x^2)T_n'(x)\\ &= -(1 - x\lambda_\nu)T_n'(x).\end{aligned}$$

In the same way, we obtain at the zeros of $T_n(x)$,

$$T_n'(x) = -T_n'(x)\sum_{0}^{n} \frac{T_n(\lambda_\nu)}{\psi'(\lambda_\nu)} \frac{1 - x\lambda_\nu}{(x - \lambda_\nu)^2}$$

and since $T_n(\lambda_\nu)$ and $\psi'(\lambda_\nu)$ are of opposite sign this gives

$$T_n'(x) = T_n'(x)\sum_{0}^{n} \left|\frac{1}{\psi'(\lambda_\nu)}\right| \frac{1 - x\lambda_\nu}{(x - \lambda_\nu)^2}. \tag{18}$$

Now $|f(\lambda_\nu)| \leqq 1$, so comparing (17) and (18) we obtain

$$|f'(x)| \leqq |T_n'(x)| \quad \textit{wherever} \quad T_n(x) = 0.$$

The equality occurs only if

$$f(\lambda_\nu) = \gamma T_n(\lambda_\nu); \qquad \nu = 0, 1, 2, \cdots, n,$$

where γ is a constant, $|\gamma| = 1$. This proves the lemma.

THEOREM II. *If $f(z)$ is a polynomial of degree n with real coefficients and if*

$$|f(\cos(\nu\pi/n))| \leqq 1, \qquad \nu = 0, 1, 2, \cdots, n,$$

then for $p = 1, 2, 3, \cdots, n$

$$|f^{(p)}(x + iy)| \leqq |T_n^{(p)}(1 + iy)|; \qquad -1 \leqq x \leqq 1, \ -\infty < y < \infty.$$

The equality occurs only if $f(z) \equiv \pm T_n(z)$.

Proof. Suppose $f(z)$ is not identically $\pm T_n(z)$. Then by Lemma IV there is a constant c, $c>1$, such that $|cf'(x)| \leqq |T_n'(x)|$ at the zeros of $T_n(x)$. Applying Theorem I to $cf(x)$ we find

$$|f^{(p)}(x+iy)| \leqq \frac{1}{c}|T_n^{(p)}(1+iy)|.$$

The case $f(x) \equiv \pm T_n(x)$ is also easily discussed.

In the above theorem in which an estimate of $f^{(p)}(z)$ was obtained throughout a strip of the complex plane the restriction that $f(z)$ have real coefficients was essential. However for points on the real axis the same bound for $|f^{(p)}(z)|$ holds even if $f(z)$ is allowed to have complex coefficients.

THEOREM III. *Let $f(z)$ be a polynomial of degree n satisfying $|f(\cos(\nu\pi/n))| \leqq 1$; $\nu=0, 1, 2, \cdots, n$. Then for $p=1, 2, 3, \cdots, n$,*

$$|f^{(p)}(x)| \leqq \frac{n^2(n^2-1)(n^2-2^2)\cdots(n^2-(p-1)^2)}{1\cdot 3\cdot 5\cdots(2p-1)}. \tag{19}$$

The equality occurs only if $f(z) \equiv \gamma T_n(z)$, $|\gamma|=1$.

Proof. If $f(z) \not\equiv \gamma T_n(z)$ then by Lemma IV there is a constant c, $c>1$, such that $|cf'(z)| \leqq |T_n'(x)|$ at the zeros of $T_n(x)$. From the Remark to Theorem I it follows that

$$|cf^{(p)}(x)| \leqq |T_n^{(p)}(1)|, \qquad -1 \leqq x \leqq 1,$$

for $p=1, 2, 3, \cdots, n$. To complete the proof of Theorem III we need only calculate $T_n^{(p)}(1)$.

Differentiating (11) p times we obtain

$$(1-z^2)T_n^{(p+2)}(z) - (2p+1)zT_n^{(p+1)}(z) + (n^2-p^2)T_n^{(p)}(z) = 0$$

from which

$$(2p+1)T_n^{(p+1)}(1) = (n^2-p^2)T_n^{(p)}(1)$$

follows. Now $T_n(1)=1$ so, using induction, we find that $T_n^{(p)}(1)$ is equal to the right-hand side of (19).

In Theorem III we have proved that the conclusions of Markoffs' theorem are true under the lighter hypothesis that $f(x)$ is bounded by 1 only at the $n+1$ points $x=\cos(\nu\pi/n)$; $\nu=0, 1, 2, \cdots, n$. This raises the question if there are $n+1$ other points in the interval $(-1, 1)$ with the same property. The answer is in the negative, however, for *if E is any closed set of points in $(-1, 1)$ which does not include all the points $x=\cos(\nu\pi/n)$ then there is a polynomial of degree n which is bounded by 1 in E, whose derivatives do not satisfy* (19). To show this suppose that E does not include the point $x=\cos(\lambda\pi/n)$. If

$\lambda=0$ or n the set E is contained in an interval of length less than 2, so an example is easily constructed. Suppose $\lambda\neq 0, n$ and let

$$P(x)=\frac{(x^2-1)T_n'(x)}{x-\cos \lambda\pi/n}, \qquad f(x)=T_n(x)+\alpha P(x), \qquad \alpha>0.$$

If ϵ is a small positive number the set E is contained in the two intervals $(-1, \cos \lambda\pi/n-\epsilon)$ and $(\cos \lambda\pi/n+\epsilon, 1)$. Call this the set E'. It is clear that if α is sufficiently small we shall have $|f(x)|\leqq 1$ in a neighborhood of the end points of the two intervals of E'. Then the maximum of $|f(x)|$ for $x\in E'$ (and this maximum is greater than 1) must occur in a small neighborhood of some point $\cos k\pi/n$ where $k\neq 0, \lambda, n$. At this maximum, say x_1, we have

$$f'(x_1)=T_n'(x_1)+\alpha P'(x_1)=0.$$

But

$$|P'(x)|\leqq A_1, \qquad -1\leqq x\leqq 1,$$

and in a neighborhood of $\cos (k\pi/n)$

$$|T_n'(x)|\geqq A_2|x-\cos (k\pi/n)|.$$

Then

$$A_2|x_1-\cos (k\pi/n)|\leqq |T_n'(x_1)|=\alpha|P'(x_1)|\leqq \alpha A_1.$$

Thus the maximum of $f(x)$ occurs within a distance $\alpha A_1/A_2=A_3\alpha$ of some point $\cos (k\pi/n)$ where $k\neq 0, \lambda, n$ and $A_3>0$ is independent of α. Hence we have by the mean value theorem

$$T_n(x_1)=T_n(\cos (k\pi/n))+\tfrac{1}{2}(x_1-\cos (k\pi/n))^2T_n''(x_2),$$
$$P(x_1)=(x_1-\cos (k\pi/n))P'(x_3)$$

where $|x_2-x_1|<|x_1-\cos (k\pi/n)|$, $|x_3-x_1|<|x_1-\cos (k\pi/n)|$. Then if

$$|T_n''(x)|\leqq A_4, \qquad -1\leqq x\leqq 1,$$

we have

$$|f(x_1)|\leqq 1+\tfrac{1}{2}A_4A_3^2\alpha^2+A_1A_3\alpha^2=1+A_5\alpha^2.$$

Thus for all small α the polynomial

$$F(x)=f(x)(1-\alpha^{3/2})$$

satisfies

$$|F(x)|<(1+A_5\alpha^2)(1-\alpha^{3/2})<1, \qquad x\in E',$$
$$|F^{(p)}(1)|=(T_n^{(p)}(1)+\alpha P^{(p)}(1))(1-\alpha^{3/2})>T_n^{(p)}(1), \quad p=1, 2, 3, \cdots, n.$$

IV. Application to entire functions

Professor Szegö has pointed out to us that by a limiting process the previous results may be applied to entire functions.

THEOREM IV. *Let $f(z)$ be an entire function, real for real z, which satisfies*
(a) $|f(z)| = o(|z|^{1/2} \exp \pi |z|^{1/2})$ *uniformly as* $|z| \to \infty$,
(b) $|f(n^2)| \leqq 1, n = 0, 1, 2, 3, \cdots$.
Then for $p = 1, 2, 3, \cdots$ *and* $x \geqq 0$

$$|f^{(p)}(x + iy)| \leqq \left| \frac{d^p}{dy^p} \cos \pi (iy)^{1/2} \right|.$$

Proof. It is evident that $f(z^2)$ is an entire function of exponential type and is $o(|z| e^{\pi|z|})$ as $|z| \to \infty$. According to a result of Valiron we may write[5]

$$f(z^2) = z\pi^{-1} \sin \pi z \left\{ \frac{f(0)}{z^2} + \sum_{-\infty}^{\infty}{}' \frac{(-1)^\nu f(\nu^2)}{\nu(z - \nu)} \right\}.$$

Then

$$f(z) = (z)^{1/2}\pi^{-1} \sin \pi (z)^{1/2} \left\{ \frac{f(0)}{z} + 2 \sum_{\nu=1}^{\infty} \frac{(-1)^\nu f(\nu^2)}{z - \nu^2} \right\}.$$

Let

$$\phi_n(z) = z(z - 2) T_n'(1 - z) = c \prod_{\nu=0}^{n} (z - \lambda_{\nu,n})$$

where $\lambda_{\nu+1,n} - \lambda_{\nu,n} > 0$ and $c \neq 0$. By Theorem II, if $-1 \leqq \delta \leqq 1$ and $0 \leqq x \leqq 2$,

$$\left| \frac{d^p}{dz^p} \sum_{\nu=0}^{k} \frac{\delta_\nu \phi_n(z)}{\phi_n'(\lambda_{\nu,n})(z - \lambda_{\nu,n})} \right| \leqq \left| \frac{d^p}{dy^p} T_n(1 - iy) \right|.$$

Here k is any integer less than or equal to n. By a change of variable we obtain, if $0 \leqq x \leqq 4n^2\pi^{-2}$, $-\infty < y < \infty$,

$$\left| \frac{d^p}{dz^p} \sum_{\nu=0}^{k} \frac{\delta_\nu \phi_n(\frac{1}{2}\pi^2 n^{-2} z)}{\phi'(\lambda_{\nu,n})(\frac{1}{2}\pi^2 n^{-2} z - \lambda_{\nu,n})} \right| \leqq \left| \frac{d^p}{dy^p} T_n(1 - \tfrac{1}{2} i \pi^2 n^{-2} y) \right|.$$

Keeping k, z, and ν fixed we allow n to increase and obtain

$$\lim_{n \to \infty} T_n(1 - \tfrac{1}{2}\pi^2 n^{-2} z) = \cos \pi (z)^{1/2},$$

$$\lim_{n \to \infty} \phi_n(\tfrac{1}{2}\pi^2 n^{-2} z) = \pi (z)^{1/2} \sin \pi (z)^{1/2}, \qquad \lim_{n \to \infty} 2\pi^{-2} n^2 \lambda_{\nu,n} = \nu^2.$$

[5] G. Valiron, *Sur la formule d'interpolation de Lagrange*, Bulletin des Sciences Mathématiques, vol. 49 (1925), pp. 203–224. Compare G. Pólya and G. Szegö, loc. cit., vol. 1, p. 117, Problem 165.

Also

$$\tfrac{1}{2}\pi^2 n^{-2}\phi_n'(\lambda_{\nu,n}) = \begin{cases} \pi^2 & \text{if } \nu = 0, \\ \tfrac{1}{2}(-1)^\nu \pi^2 & \text{if } 0 < \nu < n. \end{cases}$$

Thus we obtain

$$\left| \frac{d^p}{dz^p} \frac{(z)^{1/2} \sin \pi(z)^{1/2}}{\pi} \left\{ \frac{\delta_0}{z} + 2\sum_{\nu=1}^{k} \frac{\delta_\nu(-1)^\nu}{z-\nu^2} \right\} \right| \leqq \left| \frac{d^p}{dy^p} \cos \pi(iy)^{1/2} \right|$$

for all δ_ν which satisfy $-1 \leqq \delta_\nu \leqq 1$. Now k is an arbitrary positive integer, so the theorem follows.

We might say that Theorem IV is imperfect because in the passage to the limit from polynomials to entire functions we have lost touch with the conditions of equality. These could be discussed by reformulating Theorem I to make it apply to entire functions. A similar limiting process yields an analogue of Theorem III.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.
STANFORD UNIVERSITY,
STANFORD UNIVERSITY, CALIF.

ON DIFFERENTIAL GEOMETRY INTRINSICALLY CONNECTED WITH A SURFACE ELEMENT OF PROJECTIVE ARC LENGTH

BY

P. O. BELL

1. **Introduction.** Grove [1] has proved the existence of a remarkable class of congruences covariantly related to a surface, each congruence of which is a suitable projective substitute for the *normal congruence*. He has also shown that there is uniquely associated with a general one of these congruences a *canonical form* for the differential equations of the surface. The line of a general one of these congruences which passes through a point of the surface he has called the *R-conjugate line* at the point. The *projective normal* of Fubini is a special case of the *R*-conjugate line, and the associated canonical form is *Fubini's canonical form*.

The author [1] has presented two geometric characterizations for the *R*-conjugate line at a point of the surface, one of which exhibits this line as the cusp axis of the point with respect to the extremal curves of the integral invariant $I=\int(2Rv')^{1/2}du$. These extremal curves were characterized geometrically, but a geometric interpretation was not obtained for the intrinsic relation which was known to exist between the integral invariant I and the arc along which it is extended. The element defined by $ds=(2Rv')^{1/2}du$ is a *generalization of Fubini's element of projective arc length*. In the present paper a geometric interpretation of this general element is developed and used as a basis for the formulation of such an interpretation of its integral I. These interpretations form a geometric foundation for a detailed investigation of the properties of the surface which are intrinsically connected with a *generalized element of projective arc length*. Such an investigation is initiated in this paper.

The method employed here makes extensive use of intrinsic differentiation, differentiation along a curve with respect to projective arc length of the curve. This method consists of: first, the use of intrinsic differentiation to obtain expressions for geometric concepts, and secondly, the application of a canonical system of differential equations to reinterpret these expressions.

A euclidean geometric characterization has been obtained by Abramescu [1] for the *tangents of Darboux* at a point of a surface. In the present paper a projective generalization of an extension of this characterization is obtained. This results from the removal of certain hypothetical restrictions which have

---

Presented to the Society, December 30, 1940; §6 and §7 contain generalizations of results presented September 7, 1939, under the title *Projective invariants of a curve on a surface*; received by the editors October 4, 1940.

been placed on the base curves of the characterization and the replacement of the euclidean element of arc length by the generalized element of projective arc length.

In §5 the *associate conjugate net* is defined and a characteristic geometric property is derived for it. Similar characteristic properties are obtained for *conjugate tangents* of a surface, the *tangents of Darboux*, and the *tangents of Segre*.

An important sequence of covariant points is geometrically determined in §6. The first three points of this sequence are used in §7 to obtain geometric characterizations for the following differential invariants of a general curve of an arbitrary family on a surface: a *generalization of Fubini's asymptotic curvature*, a *generalization of Bompiani's projective curvature*, and a *projective torsion* which is introduced in this paper.

In the last section applications are made to three important special R-conjugate lines. A unique "*metric*" *invariant integral* and a corresponding *canonical form* for the differential equations of the surface are found to be associated with the *normal*. The *normal* to the surface at a point is, finally, characterized as the *cusp axis of the point with respect to the extremal curves of the associated integral*.

2. **The fundamental canonical form.** If the asymptotic net on a surface S in ordinary space is chosen as parametric, the four homogeneous coordinates $y^{(i)}$, $(i=1, 2, 3, 4)$, of a general point y of S may be taken to be solutions of differential equations in Wilczynski's (so-called) canonical form

$$y_{uu} + 2by_v + fy = 0, \qquad y_{vv} + 2a'y_u + gy = 0. \tag{2.1}$$

Consider the parametric vector equations

$$\begin{gathered} y = y(u, v), \qquad \rho = y_u - \beta y, \qquad \sigma = y_v - \alpha y, \\ \tau = y_{uv} - \alpha y_u - \beta y_v + \alpha\beta y \end{gathered} \tag{2.2}$$

where α and β are arbitrary analytic functions of u and v. Equations (2.2) define general homogeneous coordinates of four points which we denote by y, ρ, σ and τ when no confusion can arise. The line l joining the points ρ, σ is, according to Wilczynski's classification(1), an arbitrary line of the first kind and generates a congruence of the first kind as y moves over S. The reciprocal with respect to S at y of the line l is the line l' which joins the points y, τ. The line l' and congruence generated by l' as y moves over S are said to be of the second kind. If the functions α, β are suitably chosen, the points y, ρ, σ and τ become covariant points and the congruences generated by l and l' become covariant congruences.

The most general transformation of independent variables u, v which

(1) See Wilczynski [2, p. 95]. Green [1, p. 114] used this means of classifying his canonical edges of the first and second kinds.

transforms u-curves into $\bar{u}$-curves and v-curves into $\bar{v}$-curves is

$$\bar{u} = \bar{u}(u), \qquad \bar{v} = \bar{v}(v). \tag{2.3}$$

If (2.3) is performed on system (2.1), a new system is obtained whose coefficients are

$$\begin{aligned} \bar{a} &= \bar{u}_{uu}/2\bar{u}_u^2, & \bar{b} &= b\bar{v}_v/\bar{u}_u^2, & \bar{f} &= f/\bar{u}_u^2, \\ \bar{a}' &= a'\bar{u}_u/\bar{v}_v^2, & \bar{b}' &= \bar{v}_{vv}/2\bar{v}_v^2, & \bar{g} &= g/\bar{v}_v^2, \end{aligned} \tag{2.4}$$

wherein $2\bar{a}$ is the coefficient of $y_{\bar{u}}$ in the first equation and $2\bar{b}'$ is the coefficient of $y_{\bar{v}}$ in the second equation of the system.

The transformation

$$y = \mu\bar{y}, \tag{2.5}$$

where μ is an analytic function of u, v, leaves the surface undisturbed and maintains the parametric character of the asymptotic net.

Consider with Grove [1, p. 583] an invariant R of the surface—in our notation a function of the coefficients of (2.1) which is an invariant of (2.5)—with the following properties: (a) $R \neq 0$, (b) the transform $\overline{R}$ of R by (2.3) is $\overline{R} = R/\bar{u}_u\bar{v}_v$. The transform of $dudv$ by (2.3) is given by $d\bar{u}d\bar{v} = \bar{u}_u\bar{v}_v dudv$. Hence the quadratic form

$$(ds)^2 = 2Rdudv \tag{2.6}$$

is absolutely invariant under (2.3) and (2.5).

The most general transformation which preserves the parametric character of the asymptotic net and at the same time maintains Wilczynski's canonical form (2.1) is(2)

$$y = c\bar{y}/(\bar{u}_u\bar{v}_v)^{1/2}, \qquad \bar{u} = \bar{u}(u), \qquad \bar{v} = \bar{v}(v), \qquad c = \text{const.} \tag{2.7}$$

Consider the points x_u, x_v where x is defined by $x = R^h y$, $h = \text{const}$. From property (b) and equations (2.7) we obtain

$$(R^h y)_u = \bar{u}_u([\bar{u}_u\bar{v}_v\overline{R}]^h\bar{y}/[\bar{u}_u\bar{v}_v]^{1/2})_{\bar{u}}, \quad (R^h y)_v = \bar{v}_v([\bar{u}_u\bar{v}_v\overline{R}]^h\bar{y}/[\bar{u}_u\bar{v}_v]^{1/2})_{\bar{v}}. \tag{2.8}$$

These relations show clearly that x_u, x_v are covariant points if and only if $h = 1/2$. The transformation $y = x/R^{1/2}$ transforms system (2.1) into the general canonical form of Grove:

$$\begin{aligned} x_{uu} &= px + \theta_u x_u + \beta x_v, \\ x_{vv} &= qx + \gamma x_u + \theta_v x_v, \end{aligned} \tag{2.9}$$

wherein $\theta = \log R$, $\beta = -2b$, $\gamma = -2a'$, $p = -f + b\theta_v + \theta_{uu}/2 - \theta_u^2/4$ and

(2) See Wilczynski [1, p. 256].

$q=-g+a'\theta_u+\theta_{vv}/2-\theta_v^2/4$. The points x_u, x_v, x_{uv} are covariant points. The points x_u, x_v are the intersections of the *R-harmonic line* with the asymptotic u- and v-tangents to S at x, and the point x_{uv} lies on the *R-conjugate line* and is characterized geometrically in an obvious manner. The R-conjugate and R-harmonic lines are not only reciprocals with respect to S at x but as x varies over S they generate congruences which are respectively *conjugate* and *harmonic* to S.

The forms (2.6) and (2.9) are generalizations of Fubini's normal form and canonical forms respectively, and the R-conjugate line is a generalization of Fubini's projective normal. Each of these forms of Fubini corresponds, as does also the projective normal, to the selection $R=\beta\gamma$. Because of its unique association with the R-conjugate line, it seems appropriate to call the form (2.9) the *R-canonical form.*

3. **The geometric characterization of the integral invariant.** Consider any two covariant (geometrically determined) points ω_1 and ω_2 which are collinear with x but do not lie in the tangent plane to S at x. The general coordinates for ω_1 and ω_2 are given by $\omega_1=x_{uv}-cx_u-dx_v+rx$, $\omega_2=\omega_1-Rx/2$, where c, d, r and R are functions of u, v whose values depend on the selection of ω_1, ω_2. Referred to a local tetrahedron whose vertices are the points

$$x, \quad \rho = x_u - dx, \quad \sigma = x_v - cx, \quad \omega_1,$$

the equation for a general quadric of Darboux may be easily found to be

$$x_2x_3 - x_1x_4 + kx_4^2 = 0, \tag{3.1}$$

where k is an arbitrary function of u, v. The unique quadric which passes through the point ω_2 has the equation

$$2(x_2x_3 - x_1x_4) - Rx_4^2 = 0. \tag{3.2}$$

Putting $x_1=0$, we obtain the equation

$$2x_2x_3 - Rx_4^2 = 0 \tag{3.3}$$

for the cone whose vertex is the point x and which passes through the conic of intersection of the plane $x_1=0$ with the quadric (3.2). An arbitrary generator l_h' of this cone joins the points x, $2(\omega_1h+R\rho)+h^2\sigma$, $h=$arbitrary const. The reciprocal of l_h', which we denote by l_h, joins the points $\rho+hx/2$, $\sigma+Rx/h$. The equations for l_h are therefore

$$2hx_1 - h^2x_2 - 2Rx_3 = 0, \qquad x_4 = 0. \tag{3.4}$$

As h is varied while u, v are held constant the line l_h' generates the cone (3.3), and the line l_h envelops a conic $\mathcal{C}$ whose equation, when referred to the tri-

angle of reference whose vertices are the points x, $\rho = x_u - dx$, $\sigma = x_v - \iota x$ may easily be found to be

$$x_1^2 - 2Rx_2x_3 = 0. \tag{3.5}$$

This conic touches the asymptotic u- and v-tangents to S at x in the points ρ and σ, in which the reciprocal of the line joining ω_1, ω_2 intersects these tangents. This important conic plays a basic role in the geometric determinations of the present paper.

An arbitrary one-parameter family of curves on S is defined by the curvilinear equation

$$dv - \lambda(u, v)du = 0, \tag{3.6}$$

where λ is an arbitrary function of u, v. We shall denote by F_λ the family defined by (3.6) and by C_λ the curve of the family which passes through the point x.

The differential form

$$ds = (2R\lambda)^{1/2}du, \tag{3.7}$$

which corresponds to an arbitrary selection of covariant points ω_1, ω_2, is an absolute invariant under the most general transformation of independent and dependent variables maintaining the asymptotic net as parametric.

To provide a geometric interpretation for ds let Y denote a point on the tangent to C_λ at x and let N and $\overline{N}$ denote the points in which this tangent intersects the conic $\mathcal{C}$. We define the *non-euclidean distance* from x to Y to be

$$D_{xY} = (x, \overline{N}, Y, N)/2. \tag{3.8}$$

The general coordinates of N and $\overline{N}$ are given by $N = \rho + \lambda\sigma + (2R\lambda)^{1/2}x$, $\overline{N} = \rho + \lambda\sigma - (2R\lambda)^{1/2}x$. Let X denote a point near to x on the curve C_λ, where the curvilinear coordinates of x and X are (u, v) and $(u + \delta u, v + \delta v)$, respectively. Since $X = x(u + \delta u, v + \delta v)$ and the limit of $\delta v/\delta u$ as δu tends to zero is $\lambda(u, v)$, for sufficiently small values of δu the general coordinates of X may be given by the expansion

$$X = x + (x_u + \lambda x_v)\delta u + \text{terms of order } \delta u^2 \text{ at least},$$

wherein $x = x(u, v)$. Hence coordinates of X differ only by terms of order δu^2 from those of the point X_1 on the tangent to C_λ at x given by $X_1 = x + (x_u + \lambda x_v)\delta u$. Therefore the principal parts of the infinitesimal cross ratios

$$(x, \overline{N}, X, N), \qquad (x, \overline{N}, X_1, N)$$

are identical. It may be easily shown that *this principal part is the absolute differential invariant* (3.7).

To interpret geometrically the integral of the form (3.7) extended over a

finite arc of C_λ let A and B denote the end points of the arc, and let (u_0, v_0) and (u, v) denote the curvilinear coordinates of A and B, respectively. Let ϵ be a positive number, and divide the arc AB by means of intermediate points X_i, $(i=1, 2, \cdots, n-1)$, into n smaller arcs. Let the curvilinear coordinates of X_k be (u_k, v_k), where $u_n=u$ and $v_n=v$, and where

$$[(u_k - u_{k-1})^2 + (v_k - v_{k-1})^2]^{1/2} \leqq \epsilon, \qquad k = 1, 2, \cdots, n,$$

with ϵ tending to zero as n increases without limit. Then if we put $u_k - u_{k-1} = \delta u_k$ and $v_k - v_{k-1} = \delta v_k$, we have

$$I = \int_{AB} ds = \lim_{n\to\infty} \sum_{k=1}^{n} D_{X_{k-1}X_k}.$$

We have therefore

THEOREM 1. *The integral of the form* (3.7) *extended over an arc AB is the limit of the sum of infinitesimal non-euclidean distances, each of which is defined at a separate point X_{k-1} of the arc as one-half of the principal part of the corresponding cross ratio $(X_{k-1}, \overline{N}, X_k, N)$ which is geometrically determined at X_{k-1}.*

We shall call the integral I extended over an arc of S *projective arc length with respect to the form* (3.7).

It may be observed that *Fubini's element of projective arc length*

$$ds = (2\beta\gamma\lambda)^{1/2}du \tag{3.9}$$

is a special case of (3.7).

4. **The tangents of Darboux.** Let C be a curve of normal section of a surface S (in euclidean space) through a point x on S. Let M_1, M_2 be points at distances s and $-s$ from x measured along C. The line M_1M_2 intersects the tangent to C at x in a point T_a. Abramescu [1] has recently obtained the following results: *The limit of T_a as s tends to zero is a point T distinct from x. The locus of T as C varies through the set of normal sections is a rational cubic curve. The lines joining the point x to the points of inflection of this cubic curve are the tangents of Darboux.*

These results may be extended as follows. Let the curve of normal section of S be replaced by an arbitrary curve C_λ of S. Let $\overline{X}_1$, $\overline{X}_2$ denote the points at distances s and $-s$ measured from x along C_λ. Let T_s denote the point in which the tangent plane to S at x intersects the line joining $\overline{X}_1$, $\overline{X}_2$. It may be easily shown that *if the curves C and C_λ have at x a common osculating plane, the limit points of T_a and T_s as s tends to zero coincide in the point T. Thus the normal sections of S may be replaced in the characterization by less restricted curves.*

These results are capable of projective generalization as follows. Let projective arc lengths s, $-s$ be measured with respect to the form (3.7) from the point x along the curve C_λ. Let the end points of this arc be denoted by $x(s)$,

$x(-s)$. The line joining $x(s)$, $x(-s)$ intersects the tangent plane to S at x in a point which will be denoted by P_s.

THEOREM 4.1. *The limit of the point P_s as s tends to zero is a point P_0, distinct from x, which lies on the tangent to C_λ at x. The locus of P_0 as the direction of C_λ is varied at x is a rational cubic curve. The lines joining the point x to the points of inflection of this cubic are the tangents of Darboux. The line joining the three points of inflection is the R-harmonic line (the reciprocal of the R-conjugate line).*

To prove this theorem we shall need the expressions for the first, second and third derivatives of x with respect to s, where ds is defined by (3.7). By making use of the fundamental differential equations in the R-canonical form (2.4) of Grove we obtain

$$dx/ds = (x_u + \lambda x_v)/(2R\lambda)^{1/2}, \tag{4.1}$$

$$d^2x/ds^2 = (x_{uv} - cx_u - dx_v + nx)/R \tag{4.2}$$

wherein c, d and n are defined by

$$\begin{aligned} c &= (\lambda' + \theta_v\lambda^2 - \theta_u\lambda - 2\gamma\lambda^3)/4\lambda^2, \quad d = -(\lambda' + 2\beta - \theta_u\lambda + \theta_v\lambda^2)/4\lambda, \\ n &= (p + q\lambda^2)/2\lambda, \qquad\qquad \lambda' = \lambda_u + \lambda\lambda_v, \end{aligned} \tag{4.3}$$

$$d^3x/ds^3 = (Ax_{uv} + Bx_u + Cx_v + Dx)/R(2R\lambda)^{1/2} \tag{4.4}$$

wherein

$$\begin{aligned} A &= (\gamma\lambda^3 + \beta)/2\lambda, \\ B &= \theta_{uv} + \beta\gamma + \lambda\chi - c_u - c_v\lambda - b\gamma\lambda + n + c\theta_v\lambda, \\ C &= \lambda\theta_{uv} + \lambda\beta\gamma + \Pi - d_u - d_v\lambda - c\beta + n\lambda + d\theta_u, \\ D &= p_v + \beta q + n_u - cp - n\theta_u + \lambda(q_u + \gamma p + n_v - dq - n\theta_v), \\ \chi &= q + \gamma\theta_u + \gamma_u, \qquad \Pi = p + \beta\theta_v + \beta_v, \qquad \theta = \log R. \end{aligned}$$

If we take s as the independent variable, for sufficiently small values of s the general coordinates of $x(s)$ and $x(-s)$ are given by the expansions

$$x(s) = x + (dx/ds)s + (d^2x/ds^2)s^2/2 + (d^3x/ds^3)s^3/6 + \cdots, \tag{4.5}$$

$$x(-s) = x - (dx/ds)s + (d^2x/ds^2)s^2/2 - (d^3x/ds^3)s^3/6 + \cdots. \tag{4.6}$$

The general coordinates of the point P_s may be found by forming a linear combination of $x(s)$ and $x(-s)$ which is free from terms involving x_{uv}. By making use of the equations (4.1), (4.2), (4.3) and (4.4) we find that the coefficient of x_{uv} in $x(s)$ is given by $s^2/2R + A\lambda s^3/3(2R\lambda)^{3/2}$ + terms of order s^4. Since the expansion for $x(-s)$ differs from that of $x(s)$ only in the signs of its terms of odd degree in s, the coefficient of x_{uv} in $x(-s)$ must be given by

$s^2/2R - A\lambda s^3/3(2R\lambda)^{3/2}$ + terms of order s^4. Hence the point P_s, except for terms of order s^5, is represented by the linear combination

$$[1/2R + A\lambda s/3(2R\lambda)^{3/2}]x(-s) - [1/2R - A\lambda s/3(2R\lambda)^{3/2}]x(s).$$

Performing the operations indicated by this combination and omitting an unessential factor s, we find the general homogeneous coordinates of P_s to be defined by

$$2A\lambda x/3(2R\lambda)^{3/2} - x'/R + [A\lambda x''/3(2R\lambda)^{3/2} - x'''/6R]s^2 + \text{terms of order } s^3 \text{ at least.} \tag{4.7}$$

The limit of the point P_s as s tends to zero is the point P_0, distinct from x, which lies on the tangent to C_λ at x and whose general homogeneous coordinates (except for a non-essential factor) are given by

$$(\gamma\lambda^3 + \beta)x - 6\lambda x_u - 6\lambda^2 x_v. \tag{4.8}$$

As the direction λ is varied while u, v are held constant, the point P_0 describes the rational cubic curve whose equation is

$$\gamma x_3^3 + \beta x_2^3 + 6x_1x_2x_3 = 0, \tag{4.9}$$

when referred to the triangle whose vertices are the points x, x_u and x_v. The three points of inflection lie on the line whose equation is $x_1=0$. Since the points x_u, x_v also lie on this line, it is the R-harmonic line. The equation

$$\gamma x_3^3 + \beta x_2^3 = 0 \tag{4.10}$$

represents the three lines which join the point x to the different points of inflection. The form of equation (4.10) identifies these lines as the tangents of Darboux. This completes the proof of the results stated in the opening paragraph of this section.

If λ satisfies the equation $\gamma\lambda^3+\beta=0$ it is a direction of Darboux. Therefore,

THEOREM 4.2. *The point P_0 which corresponds to a direction λ lies on the R-harmonic line if and only if λ is a direction of Darboux.*

Since the coefficient of x_{uv} in the form (4.4) for the coordinates of the point d^3x/ds^3 is $(\gamma\lambda^3+\beta)/(2R\lambda)^{3/2}$, we have

THEOREM 4.3. *The point d^3x/ds^3, where $ds=(2R\lambda)^{1/2}du$, $\lambda\neq\infty$, and R is an arbitrary function of (u, v), lies in the tangent plane to S at x if and only if the direction λ is a direction of Darboux.*

5. **Theorems on nets.** The curvilinear differential equation of an arbitrary net N_{λ,λ_1} on the surface S can be written in the form

$$(5.1) \qquad (dv - \lambda du)(dv - \lambda_1 du) = 0,$$

where λ, λ_1 are arbitrary functions of u, v. There exists on S a conjugate net uniquely determined by the net N_{λ,λ_1}, such that at each point of S its tangents separate the tangents of the net N_{λ,λ_1} harmonically. This net will be called the *associate conjugate net of the net* N_{λ,λ_1} (Green [2, p. 213]). Its differential equation is found to be

$$(5.2) \qquad dv^2 - \lambda\lambda_1 du^2 = 0.$$

A remarkable characteristic property of this net may be obtained as follows. Let projective arc lengths s, $-s$ be measured with respect to the form $ds = (2Rv')^{1/2}du$ from the point x along the curves C_λ, C_{λ_1}. Let the end points of these arcs be denoted by $x(s)$, $x(-s)$ and $x_1(s)$, $x_1(-s)$, respectively. The line joining the points $x(s)$, $x_1(s)$ and the line joining the points $x(s)$, $x_1(-s)$ intersect the tangent plane to S at x in points which we denote by Q_s and $\overline{Q}_s$, respectively.

THEOREM 5.1. *The points Q_s and $\overline{Q}_s$ approach limit points Q_0 and $\overline{Q}_0$, respectively, as s tends to zero. The tangents to S at x joining x to the separate points Q_0 and $\overline{Q}_0$ are the tangents at x of the associate conjugate net of the net N_{λ,λ_1}.*

To prove this theorem consider the general coordinates of the points $x(s)$, $x_1(s)$, $x_1(-s)$. The coordinates of $x(s)$ for small values of s are given by the expansion (4.5), wherein the derivatives are defined by (4.1), (4.2) and (4.4). The coordinates of $x_1(s)$ are obtained from those of $x(s)$ by replacing λ by λ_1. The coordinates of $x_1(-s)$ may be obtained from those of $x_1(s)$ by changing the signs of the terms containing odd exponents of s. The coefficient of x_{uv} in $x(s)$ is given by $s^2/2R + (\gamma\lambda^3+\beta)s^3/6(2R\lambda)^{3/2}$ + terms of order s^4 at least. The coefficients of x_{uv} in $x_1(s)$ and $x_1(-s)$ are (except for terms of order s^4) given by $s^2/2R + (\gamma\lambda_1^3+\beta)s^3/6(2R\lambda_1)^{3/2}$ and $s^2/2R - (\gamma\lambda_1^3+\beta)s^3/6(2R\lambda_1)^{3/2}$, respectively. Therefore, except for terms of order s^3 at least, the points Q_s, $\overline{Q}_s$ are given by

$$(1/2R + [\gamma\lambda^3 + \beta]s/6[2R\lambda]^{3/2})x_1(s) - (1/2R + [\gamma\lambda_1^3 + \beta]s/6[2R\lambda_1]^{3/2})x(s),$$

$$(1/2R + [\gamma\lambda^3 + \beta]s/6[2R\lambda]^{3/2})x_1(-s) - (1/2R - [\gamma\lambda_1^3 + \beta]s/6[2R\lambda_1]^{3/2})x(s),$$

respectively. The developments for $x(s)$, $x_1(s)$, $x_1(-s)$ which are needed are (except for terms of order s^2 at least)

$$x(s) = x + (x_u + \lambda x_v)s/(2R\lambda)^{1/2} + \cdots,$$

$$x_1(s) = x + (x_u + \lambda_1 x_v)s/(2R\lambda_1)^{1/2} + \cdots,$$

$$x_1(-s) = x - (x_u + \lambda_1 x_v)s/(2R\lambda_1)^{1/2} + \cdots.$$

Performing the operations indicated in the expressions for Q_s, $\overline{Q}_s$ and neglecting unessential factors, we obtain the following forms:

$$(5.3)\qquad (\gamma\lambda_2 - \beta/\lambda_2^2)(\lambda + \lambda_1 + \lambda_2)x + x_u - \lambda_2 x_v,$$

$$(5.4)\qquad (\gamma\lambda_2 + \beta/\lambda_2^2)(\lambda + \lambda_1 - \lambda_2)x - x_u - \lambda_2 x_v, \qquad \lambda_2 = (\lambda\lambda_1)^{1/2},$$

for the general coordinates of Q_0 and $\overline{Q}_0$, respectively. Since the directions of the lines joining x to the points Q_0 and $\overline{Q}_0$ are, respectively, $-(\lambda\lambda_1)^{1/2}$ and $(\lambda\lambda_1)^{1/2}$, the proof of the theorem is complete.

Consider, as a special case of the net N_{λ,λ_1}, the conjugate net N_λ. If we replace λ_1 by $-\lambda$ in (5.3) and (5.4), we obtain the forms

$$(5.5)\qquad (i\gamma\lambda + \beta/\lambda^2)i\lambda x + x_u - i\lambda x_v,$$

$$(5.6)\qquad (i\gamma\lambda - \beta/\lambda^2)i\lambda x + x_u + i\lambda x_v, \qquad i = (-1)^{1/2},$$

for the coordinates of the associated points Q_0 and $\overline{Q}_0$, respectively. If, however, we replace λ by $-\lambda_1$, the forms (5.5) and (5.6) are interchanged. The reason for this rather surprising fact becomes clear if we notice the effect of these replacements in the equation

$$(5.7)\qquad ds = (2R\lambda)^{1/2}du = (2R\lambda_1)^{1/2}du_1$$

which determines at x the sense of direction of x along C_{λ_1} which corresponds to a given sense of direction of x along C_λ. If we put $-\lambda$ for λ_1 in (5.7), we have at x

$$(5.8)\qquad du = idu_1.$$

If, however, we put $-\lambda_1$ for λ in (5.7), we have at x

$$(5.9)\qquad du = -idu_1.$$

The senses of direction defined for x along C_{λ_1} by (5.8) and (5.9) are, clearly, opposed to one another. Hence, the order of the determination of the points Q_0, $\overline{Q}_0$ which corresponds to (5.9) is the reverse of that which corresponds to (5.8).

The line joining the points Q_0, $\overline{Q}_0$ has the equation

$$(5.10)\qquad x_1 + \gamma\lambda^2 x_2 + \beta x_3/\lambda^2 = 0.$$

It will be called the N_λ-*associate* of the R-harmonic line. This line intersects the R-harmonic line in a point of the tangent to S at x whose equation is

$$(5.11)\qquad \gamma\lambda^2 x_2 + \beta x_3/\lambda^2 = 0.$$

This tangent will be called the N_λ-*correspondent* of the tangent to C_λ at x.

The following theorems are immediate consequences of the form of equation (5.11):

THEOREM 5.2. *The N_λ-correspondent of the tangent to a non-asymptotic curve C_λ at x coincides with this tangent if and only if it is a tangent of Darboux.*

THEOREM 5.3. *The tangent to a non-asymptotic curve C_λ at x and its N_λ-correspondent are conjugate tangents if and only if λ is a direction of Segre.*

If λ_1 is made to approach λ, the point Q_0 tends to a limit point Q whose general homogeneous coordinates are found, by replacing λ_1 by λ in (5.3), to be given by

$$3(\gamma\lambda^3 - \beta)x + \lambda x_u - \lambda^2 x_v. \tag{5.12}$$

Since the ratio of the coefficient of x_v to that of x_u in (5.12) is equal to $-\lambda$, we have

THEOREM 5.4. *The line joining the points x and Q and the tangent to C_λ at x are conjugate tangents.*

The equation obtained by equating to zero the coefficient of x in (5.12) is satisfied by the three directions of Segre. Hence we have

THEOREM 5.5. *The tangent to C_λ at x is a tangent of Segre if and only if the associated point Q is the intersection of the R-harmonic line with the conjugate of the tangent to C_λ at x.*

As λ is varied while u, v are held constant, the point Q describes a rational cubic curve whose equation, referred to the triangle whose vertices are x, x_u, x_v, is easily found to be

$$\gamma x_3^3 + \beta x_2^3 - \tfrac{1}{3}(x_1x_2x_3) = 0. \tag{5.13}$$

This cubic has geometric properties similar to those of the cubic (4.9).

6. **A sequence of covariant points.** The points of the sequence $x^{(j)} = d^j x/ds^j$, $(j=1, 2, \cdots)$, where $ds = (2R\lambda)^{1/2}du$, are covariant. A general one of them is intrinsically and projectively related to the curve C_λ at x. A geometric interpretation of this relation will be presented in this section. The points $x^{(j)} - x^{(j-1)}$ and $x^{(j)} + x^{(j-1)}$ will play basic roles in this interpretation.

Let $C(j, -)$, $C(j, +)$, and $C(j)$ denote, respectively, the curves described by the points $x^{(j)} - x^{(j-1)}$, $x^{(j)} + x^{(j-1)}$, and $x^{(j)}$ as x moves along C_λ.

The point $x' = (x_u + \lambda x_v)/(2R\lambda)^{1/2}$ is, clearly, *the intersection of the tangent to C_λ at x with the R-harmonic line.* If the fundamental points ω_1, ω_2, $\omega_1 - \omega_2 = Rx/2$, are selected on the R-conjugate line, the associated conic $\mathcal{C}$, referred to the triangle whose vertices are x, x_u, x_v, has the equation (3.5). This conic intersects the tangent to C_λ at x in the points $x' + x$ and $x' - x$. The tangents at these points to the curves $C(1, -)$, $C(1, +)$ intersect the tangent at x' to the curve $C(1)$ in the points $x'' + x'$ and $x'' - x'$. *The point x'' is clearly the harmonic conjugate of the point x' with respect to the points $x'' + x'$ and $x'' - x'$.* Similarly the tangents at $x'' \pm x'$ to the curves $C(2, \pm)$ intersect the tangent at x'' to the curve $C(2)$ in the points $x''' + x''$ and $x''' - x''$. *The point x''' is the harmonic conjugate of x'' with respect to the points $x''' \pm x''$.* It

is quite clear that the entire sequence $x^{(j)} \pm x^{(j-1)}$, $(j=1, 2, \cdots)$, is determined in this way, such that $x^{(k)}$ *is the harmonic conjugate of* $x^{(k-1)}$ *with respect to the points* $x^{(k)}+x^{(k-1)}$ *and* $x^{(k)}-x^{(k-1)}$ *for an arbitrary integer* $j=k$.

The geometric locations of the points x'' and x''' will help to form geometric interpretations for the second and third order differential invariants of C_λ known as projective curvature and projective torsion.

7. **Differential invariants of a curve on a surface.** Fubini's asymptotic curvature and Bompiani's projective curvature are second order differential invariants of a curve C_λ on a surface S. Bompiani's projective torsion of C_λ is a third order differential invariant. These invariants have been defined analytically by the respective forms

$$\kappa_a = R(du\delta^2 v - dv\delta^2 u)/\phi_2^{3/2}, \tag{7.1}$$

$$\kappa = \kappa_a + \bar{\phi}_3/2\phi_2^{3/2}, \tag{7.2}$$

$$\begin{aligned} T = {} & [R(du\delta^3 v - dv\delta^3 u)\phi_2 - \kappa_a\phi_2^{3/2}\phi_3/2 \\ & + 3R(\beta du^2\delta^2 u - \gamma dv^2\delta^2 v)\phi_2 - \phi_3\bar{\phi}_3/4 + \bar{\psi}_2\phi_2^2 - \bar{\phi}_4\phi_2]/\phi_2^3 \end{aligned} \tag{7.3}$$

wherein $\phi_2=2Rdudv$, $\phi_3=2R(\beta du^3+\gamma dv^3)$, $\bar{\phi}_3=2R(\beta du^3-\gamma dv^3)$, $\psi_2=(p+3\beta\Psi/2)du^2-(q+3\gamma\Phi/2)dv^2$, $\bar{\phi}_4=\beta R\Phi du^4-\gamma R\Psi dv^4$, $\Psi=(\log\beta R)_v$, $\Phi=(\log\gamma R)_u$, $dv/du=\lambda$, $R=\beta\gamma$, and $\delta^2 u$, $\delta^2 v$, $\delta^3 u$, $\delta^3 v$ represent contravariant differentials constructed with respect to the form[3] ϕ_2.

Bompiani[4] has obtained geometric interpretations for the infinitesimal invariants $\kappa_a\phi_2^{1/2}$, $\kappa\phi_2^{1/2}$, and $T\phi_2^{1/2}$, where $\phi_2=2\beta\gamma dudv$. However he has given no geometric interpretation for κ_a, κ, or T.

The invariants κ_a, κ, T are special members of the classes which we introduce and denote by $\{\kappa_a\}$, $\{\kappa\}$, $\{T\}$, particular members of which may be obtained by selecting R from among all of the functions for which $(2R\lambda)^{1/2}du$ is an absolute invariant of C_λ. In this section a general member of each of the classes $\{\kappa_a\}$ and $\{\kappa\}$ will be characterized geometrically, and a general member of a new class of projective torsions $\{\overline{T}\}$ will be introduced and characterized geometrically. Each of these general invariants will be designated by the name of its corresponding original invariant together with the added phrase "*with respect to the form* $ds=(2R\lambda)^{1/2}du$." The members of the class $\{\overline{T}\}$ will be shown to be very closely related to corresponding members of the class $\{T\}$. However, a purely geometric interpretation for a general member of $\{T\}$ has not been discovered by the author.

---

(3) It should be observed that these differential invariants are not the usual ones unless $R=\beta\gamma$. If this restriction is removed they become extensions of the usual ones.

(4) See E. Bompiani, "I fondamenti geometrici della teoria proiettiva delle curve e delle superficie," Appendix II of *Geometria Proiettiva Differenziale*, G. Fubini and E. Čech [1, pp. 691–697].

Consider the conic $\mathcal{C}$ whose equation is (3.5) referred to the triangle whose vertices are x, ρ, σ. The lines which are tangent to $\mathcal{C}$ at the points in which the tangent to C_λ at x intersects $\mathcal{C}$ will be called the $\mathcal{C}_\lambda$-*associates* of the line l joining ρ, σ and will be denoted by $l_{\lambda,+}$ and $l_{\lambda,-}$. Their respective equations referred to the triangle whose vertices are x, ρ, σ are easily found to be

$$x_1 + (2R\lambda)^{1/2}x_2 + (2R/\lambda)^{1/2}x_3 = 0, \tag{7.4}$$

$$x_1 - (2R\lambda)^{1/2}x_2 - (2R/\lambda)^{1/2}x_3 = 0. \tag{7.5}$$

The crossjoins of the intersections of the lines $l_{\lambda,+}$ and $l_{\lambda,-}$ with the u- and v-tangents of S at x will be called the $\overline{\mathcal{C}}_\lambda$-*associates* of l and will be denoted by $\overline{l}_{\lambda,+}$ and $\overline{l}_{\lambda,-}$. Their respective equations are

$$x_1 - (2R\lambda)^{1/2}x_2 + (2R/\lambda)^{1/2}x_3 = 0, \tag{7.6}$$

$$x_1 + (2R\lambda)^{1/2}x_2 - (2R/\lambda)^{1/2}x_3 = 0. \tag{7.7}$$

Let l_2' denote the line joining the points x and x'' and let l_2 denote its reciprocal with respect to S at x. The R_λ-*correspondent*[8] of the tangent to C_λ at x intersects the R-harmonic line, the $\overline{\mathcal{C}}_\lambda$ associates of the R-harmonic line, and the line l_2 in the points which we denote by K_1, $\overline{K}_\pm$, K_2, respectively. Referred to the triangle whose vertices are the points x, x_u, x_v, the equations for the above named lines are, respectively, $\gamma\lambda^2x_3+\beta x_2=0$, $x_1=0$, (7.6) and (7.7), and $x_1+dx_2+cx_3=0$ where c and d are given by (4.3). The local coordinates of $K_1, \overline{K}_+, \overline{K}_-, K_2$ are easily found to be $(0, \gamma\lambda^2, -\beta)$, $([2R\lambda]^{1/2}[\gamma\lambda^2+\beta]/\lambda, \gamma\lambda^2, -\beta)$, $(-[2R\lambda]^{1/2}[\gamma\lambda^2+\beta]/\lambda, \gamma\lambda^2, -\beta)$, $(-d\gamma\lambda^2+c\beta, \gamma\lambda^2, -\beta)$, respectively. Since

$$-d\gamma\lambda^2 + c\beta = (\lambda' + \theta_v\lambda^2 - \theta_u\lambda)(\gamma\lambda^2 + \beta)/4\lambda^2,$$

the cross ratios $(x, K_1, K_\pm, K_2)$ may be found by evaluating $(\infty, 0, \pm[2R\lambda]^{1/2}, [\lambda'+\theta_v\lambda^2-\theta_u\lambda]/4\lambda)$ to be $\pm(\lambda'+\theta_v\lambda^2-\theta_u\lambda)/4\lambda(2R\lambda)^{1/2}$. But the expression for the asymptotic curvature of C_λ at x with respect to the form $(2R\lambda)^{1/2}du$ becomes

$$(\lambda' + \theta_v\lambda^2 - \theta_u\lambda)/2\lambda(2R\lambda)^{1/2}, \qquad \theta = \log R, \tag{7.8}$$

when u is regarded as the independent variable. Thus we have proved

THEOREM 7.1. *The asymptotic curvature of C_λ at x with respect to the form $ds=(2R\lambda)^{1/2}du$ is characterized by the cross ratio equations $\kappa_a=2(x, K_1, \overline{K}_+, K_2)=-2(x, K_1, \overline{K}_-, K_2)$.*

The expression for the projective curvature of C_λ at x with respect to the form $ds=(2R\lambda)^{1/2}du$ becomes

$$(\lambda' + \beta - \theta_u\lambda + \theta_v\lambda^2 - \gamma\lambda^3)/2\lambda(2R\lambda)^{1/2} \tag{7.9}$$

[8] See P. O. Bell [2, p. 392] and *The R_λ-correspondent of the tangent to an arbitrary curve of a non-ruled surface*, Bulletin of the American Mathematical Society, vol. 47 (1941), pp. 509–511.

when u is regarded as the independent variable. To characterize this invariant let M_1, $\overline{M}_{\pm}$, M_2 denote the intersections of the tangent to $C_{-\lambda}$ at x with the R-harmonic line, the $\overline{\mathcal{C}}_\lambda$-associates of the R-harmonic line, and the line l_2, respectively. The local coordinates of M_1, $\overline{M}_{\pm}$, M_2 are readily found to be $(0, 1, -\lambda)$, $(\pm 2[2R\lambda]^{1/2}, 1, -\lambda)$, $([\lambda'+\beta-\theta_u\lambda+\theta_v\lambda^2-\gamma\lambda^3]/2\lambda, 1, -\lambda)$. Hence the values of the cross ratios $(x, M_1, \overline{M}_{\pm}, M_2)$ are equal to $\pm(\lambda'+\beta-\theta_u\lambda+\theta_v\lambda^2-\gamma\lambda^3)/4\lambda(2R\lambda)^{1/2}$. Therefore

THEOREM 7.2. *The projective curvature of C_λ at x with respect to the form $ds=(2R\lambda)^{1/2}du$ is characterized by the cross ratio equations $\kappa=2(x, M_1, \overline{M}_+, M_2)=-2(x, M_1, \overline{M}_-, M_2)$.*

The curvilinear differential equation for the extremal curves of the integral invariant $\int(2R\lambda)^{1/2}du$ is easily found to be

$$\lambda' + \theta_v\lambda^2 - \theta_u\lambda = 0, \qquad \theta = \log R. \tag{7.10}$$

We shall call these curves the *projective geodesics* of the surface *with respect to the form* $ds=(2R\lambda)^{1/2}du$. The projective geodesics of Fubini are obtained by putting $R=\gamma\beta$ in (7.10).

The curvilinear differential equation for the union curves of the R-conjugate congruence is found by analogy with (7.10) to be

$$\lambda' + \beta - \theta_u\lambda + \theta_v\lambda^2 - \gamma\lambda^3 = 0, \qquad \theta = \log R. \tag{7.11}$$

If we consider the forms (7.8) and (7.9) in connection with equations (7.10) and (7.11), respectively, we have at once

THEOREM 7.3. *A curve C_λ is a projective geodesic of S with respect to a form $ds=(2R\lambda)^{1/2}du$ if and only if at each of its points its asymptotic curvature with respect to the same form vanishes.*

THEOREM 7.4. *A curve C_λ is a union curve of an R-conjugate congruence if and only if at each of its points its projective curvature with respect to the form $ds=(2R\lambda)^{1/2}du$ vanishes.*

Let $l'_{2,\pm}$ denote the reciprocals with respect to S of the $\overline{\mathcal{C}}_\lambda$-associates of the line l_2, and l'_3 denote the line joining the points x and x'''. Let $\overline{\pi}_{2,\pm}$ and π_3 denote the planes which pass through the respective lines $l'_{2,\pm}$ and l'_3 and which have the tangent to C_λ at x as common intersection. Let π and π_2 denote, respectively, the tangent plane to S at x and the osculating plane of C_λ at x. The expression defined by the cross ratio equations

$$\overline{T} = (\pi_2, \pi, \pi_3, \overline{\pi}_{2,+}) = -(\pi_2, \pi, \pi_3, \overline{\pi}_{2,-}) \tag{7.12}$$

is an absolute differential invariant of C_λ of the third order. Because of the suitability of this invariant as a projective substitute for torsion, the author

proposes that it be called the *projective torsion* of C_λ at x *with respect to the form* $ds=(2R\lambda)^{1/2}du$.

The polar reciprocals of the planes π_2, π, π_3, $\overline{\pi}_{2,\pm}$ with respect to a quadric of Darboux are the points M_2, x, M_3, $\overline{M}_{2,\pm}$ all of which lie on the tangent to $C_{-\lambda}$ at x. The points M_2, M_3 are the intersections of this tangent with the lines l_2, l_3 which are the reciprocals of l_2', l_3' with respect to a quadric of Darboux, and the points $\overline{M}_{2,\pm}$ are the intersections of this tangent with the $\overline{C}_\lambda$-associates of l_2. Since the cross ratio of four planes is equal to that of the four points which are their polar reciprocals with respect to a quadric of Darboux, the expression for $\overline{T}$ may be obtained by evaluating the cross ratio $(M_2, x, M_3, \overline{M}_{2,+})$. Making use of the local coordinates of M_2, x, M_3, $\overline{M}_{2,\pm}$, respectively, $(M/2\lambda, 1, -\lambda)$, $(\infty, 1, -\lambda)$, $(2\lambda[\Pi-\chi\lambda^2-d'+c'\lambda-c\beta+d\gamma\lambda^2+d\theta_u-c\theta_v\lambda^2]/[\gamma\lambda^3+\beta], 1, -\lambda)$, $(M/2\lambda\pm 2[2R\lambda]^{1/2}, 1, -\lambda)$, wherein $M=\lambda'+\beta-\theta_u\lambda+\theta_v\lambda^2-\gamma\lambda^3$, $\Pi=p+\beta\theta_v+\beta_v$, $\chi=q+\gamma\theta_u+\gamma_u$, $\theta=\log R$, we find

$$\overline{T} = [2\lambda(M' + 2\Pi\lambda - 2\chi\lambda^3) - M(3\lambda' + \beta + \theta_u\lambda + \theta_v\lambda^2 + \gamma\lambda^3)]/4\lambda(2R\lambda)^{1/2}(\gamma\lambda^3 + \beta). \tag{7.13}$$

The invariant T given by (7.3) with $\beta\gamma$ replaced by a function R may be shown to be related to the author's invariant $\overline{T}$ by the simple formula

$$\overline{T}d\sigma = Tds,$$

in which $d\sigma$ is the *projective linear element* defined by $d\sigma=(\beta du^3+\gamma dv^3)/dudv$ and ds is the *generalized element of projective arc length* defined by $ds=(2R\lambda)^{1/2}du$.

If at a general point x of a curve C_λ the osculating plane of C_λ contains the point x''', the osculating plane is stationary and the curve C_λ is a plane curve. In this case the planes π_2 and π_3 coincide at every point of C_λ; and $\overline{T}$, therefore, vanishes identically along C_λ. Moreover, equation (7.13) shows that if C_λ is not an asymptotic curve, $\overline{T}$ can vanish identically along C_λ only if its numerator does. If the numerator of $\overline{T}$ vanishes identically along C_λ, the points M_2, M_3 which correspond to a general point of C_λ must coincide. Hence, their polar reciprocals π_2, π_3 must also coincide and the osculating plane π_2 is, therefore, stationary. To summarize:

THEOREM 7.5. *A curve C_λ is a plane curve if and only if at a general one of its points its projective torsion with respect to the form $ds=(2R\lambda)^{1/2}du$ is equal to zero.*

The equation obtained by setting the numerator of the expression for $\overline{T}$ equal to zero is the curvilinear differential equation for all of the plane curves of S.

It may be observed that although all of the results up to this point have been developed in association with a general form $ds=(2R\lambda)^{1/2}du$, a considerable number of them are, clearly, independent of R.

8. **Special conjugate congruences.** In a recent paper [1, p. 681] the author has shown that the R-conjugate line may be characterized geometrically in association with the curves which a pair of covariant points ω_1, ω_2 describe as x moves along C_λ. To complete this characterization for special cases of the R-conjugate line, corresponding pairs of points ω_1, ω_2 must be located.

The projective normal is the special case of the R-conjugate line which may be defined in association with the points ω_1, ω_2 given by

$$\omega_1 = \tau - 2a'by, \qquad \omega_2 = \tau + 2a'by \tag{8.1}$$

wherein $\tau = y_{uv} - \alpha y_u - \beta y_v + \alpha\beta y$. Geometric determinations, which do not depend on the definitions of the quadrics of Wilczynski and Lie, have been given by the author [2, p. 392] for these points. These determinations supplement the method of [1, p. 681] to complete a simple geometric determination of the projective normal.

The following considerations serve to introduce another important special case of the R-conjugate line. Let l_k denote a general line of the *first canonical pencil*. The line l_k intersects the tangents of S at y to the asymptotic u- and v-curves in the points ρ, σ defined in (2.2) where

$$\begin{aligned} \beta &= (\log a'^2b)_u/k - (\log a'b)_u/2, \\ \alpha &= (\log b^2a')_v/k - (\log a'b)_v/2. \end{aligned} \tag{8.2}$$

The lines l_2, l_3, l_4 and l_∞ are the *first directrix of Wilczynski*, the *reciprocal of the axis of Čech*, the *first canonical edge of Green*, and the *reciprocal of the projective normal*, respectively. As y moves over S the points ρ, σ of l_k generate transversal surfaces S_ρ and S_σ of the congruence described by l_k. The v-tangent at ρ to S_ρ intersects l_k', the reciprocal of l_k, in the point which we denote by η_k whose coordinates are given by $\eta_k = \tau - \beta_v y$, where α, β are the functions (8.2). Likewise the u-tangent at σ to S_σ intersects l_k' in the point which we denote by ξ_k whose coordinates are given by $\xi_k = \tau - \alpha_u y$. If we let $\omega_1 = \eta_k$, $\omega_2 = \xi_k$, we have $\omega_1 - \omega_2 = (\log b/a')_{uv} y/k$. Hence we have

THEOREM 8.1. *If the fundamental points ω_1, ω_2 are chosen as the points η_k, ξ_k on a canonical line l_k' of the second kind, the associated R-conjugate line is independent of k and joins the points y, τ defined by (2.2) wherein the associated functions α, β are given by $\alpha = -(\log R)_v/2$, $\beta = -(\log R)_u/2$ with $R = (\log b/a')_{uv}$.*

Consider the following problem. *Find an integral of the form $\int (2R\lambda)^{1/2} du$ with the following properties*: (i) *it is absolutely invariant under the most general transformation of variables which maintains the parametric character of the asymptotic net of S and the metric properties of any orthogonal cartesian coordinate system to which S may be referred*, (ii) *the normal is the cusp axis of x with respect to the extremal curves of the integral.*

This problem is found to possess a unique solution (except for an unessential arbitrary constant factor). Let any orthogonal system of cartesian coordinates be established in ordinary metric space. Let x denote a vector from the origin to a point on the surface S wherein the coordinates $x^{(i)}$, $(i=1, 2, 3)$, are measured with respect to equal units on the three axes. Let the asymptotic net be parametric. Let X denote the unit normal to S at x. The first fundamental form is given by $Edu^2+2Fdudv+Gdv^2$ where E, F, G are defined by the scalar products

$$E = (x_u \mid x_u), \qquad F = (x_u \mid x_v), \qquad G = (x_v \mid x_v).$$

The second fundamental form is given by $2D'dudv$ where $D' = (X \mid x_{uv})$. The fundamental differential equations which are satisfied by $x^{(i)}$, $X^{(i)}$ are known to be

$$\begin{aligned} x_{uu} &= \{11, 1\}x_u + \{11, 2\}x_v, \\ x_{uv} &= \{21, 1\}x_u + \{12, 2\}x_v + D'X, \\ x_{vv} &= \{22, 1\}x_u + \{22, 2\}x_v, \\ X_u &= FD'x_u/H^2 - ED'x_v/H^2, \\ X_v &= -GD'x_u/H^2 + FD'x_v/H^2, \qquad H^2 = EG - F^2, \end{aligned} \tag{8.3}$$

where $\{ik, l\}$ denote the Christoffel symbols(6) of the second kind for the first fundamental form.

The *equation of Gauss*,

$$HK = (H\{11, 2\}/E)_v - (H\{12, 2\}/E)_u, \tag{8.4}$$

where K is the total curvature defined by $K=-D'^2/H^2$, and the *equations of Codazzi*,

$$D'_u - (\{11, 1\} - \{12, 2\})D' = 0, \quad D'_v - (\{22, 2\} - \{21, 1\})D' = 0, \tag{8.5}$$

are the three integrability conditions of (8.3).

By means of the definitions of the Christoffel symbols it is easy to show that

$$(\log H)_u = \{11, 1\} + \{12, 2\}, \qquad (\log D')_v = \{22, 2\} - \{21, 1\}.$$

Making use of these identities and the Codazzi equations (8.5) to solve for $\{11, 1\}$, $\{12, 2\}$, $\{22, 2\}$ and $\{21, 1\}$ in terms of H and D', we obtain the following form for the first three differential equations of (8.3):

$$\begin{aligned} x_{uu} &= \theta_u x_u + \{11, 2\}x_v, \qquad x_{uv} = \nu_v x_u + \nu_u x_v + D'X, \\ x_{vv} &= \{22, 1\}x_u + \theta_v x_v, \end{aligned} \tag{8.6}$$

where $\theta=\log (HD')^{1/2}$, $\nu=\log (H/D')^{1/2}$.

(6) See, e.g., E. P. Lane [1, p. 222].

The homogeneous cartesian coordinates $x^{(1)}$, $x^{(2)}$, $x^{(3)}$, 1 of x may be regarded as special projective coordinates of x which are referred to the tetrahedron of reference whose faces are the three cartesian coordinate planes and the plane at infinity. The unit point must, of course, be properly selected so that the system is cartesian. Let us determine the transformation $\mu x = \bar{x}$ which converts homogeneous cartesian coordinates x into coordinates $\bar{x}$ which satisfy the differential equations in the unique R-canonical form whose associated R-conjugate line is the normal to S at $\bar{x}$. Differentiating the equation $\bar{x} = \mu x$ we find $\bar{x}_u = \mu_u x + \mu x_u$, $\bar{x}_v = \mu_v x + \mu x_v$, $\bar{x}_{uv} = \mu_{uv} x + \mu_v x_u + \mu_u x_v + \mu x_{uv}$. Let us choose μ in such a way that $\bar{x}_{uv}$ lies on the normal to S at x. The point $x_{uv} - \nu_v x_u - \nu_u x_v$ is the point at infinity on the normal to S at x, for its fourth coordinate is, clearly, zero and the second equation of (8.6) shows that its first three coordinates are direction numbers of the normal to S at x. The coordinates of any selected point on the normal to S at x may be represented by a suitably chosen linear combination of the corresponding coordinates of the points x and $x_{uv} - \nu_v x_u - \nu_u x_v$. The point $\bar{x}_{uv}$ will therefore lie on the normal to S at x if and only if μ satisfies the equations $(\log \mu)_v = -\nu_v$, $(\log \mu)_u = -\nu_u$. Integrating these we obtain $\mu = k(D'/H)^{1/2}$, k = arbitrary constant. The effect of the transformation $\bar{x} = k(D'/H)^{1/2} x$ on the first and third equations of (8.6) is to produce a system of the form (2.4) wherein

$$\theta = \log (D'^2/H)^{1/2}, \qquad \beta = \{11, 2\}, \qquad \gamma = \{22, 1\},$$

$$p = [(H_u/H)^2 - (D'_u/D')^2]/4 + \{11, 2\}[\log H/D']_v/2 - [(H/D')^{1/2}]_{uu}/(H/D')^{1/2},$$

$$q = [(H_v/H)^2 - (D'_v/D')^2]/4 + \{22, 1\}[\log H/D']_u/2 - [(H/D')^{1/2}]_{vv}/(H/D')^{1/2}.$$

The normal is therefore the R-conjugate line for which $R = k(D'^2/H)^{1/2}$, k = arbitrary const. Since, moreover, $K = -D'^2/H^2$, we have $R = cK^{1/4}D'$, c = arbitrary const. The differential form $2K^{1/4}D'dudv$ is absolutely invariant under all transformations which maintain the orthogonal cartesian properties of the coordinate system and the parametric character of the asymptotic net. Both the total curvature K of S at x and the second fundamental form $2D'dudv$ are intrinsically connected with S at x. Let us recall that if we consider a point $x(s+\Delta s)$ near to $x(s)$, where s denotes length of arc along a curve C_λ, for an arbitrarily selected parametric net, the associated second fundamental form of S is equal in absolute value to twice the principal part of the distance from the point $x(s+\Delta s)$ to the plane which is tangent to S at $x(s)$.

The above results in combination with Theorem (2.3) [1, p. 681] yield the following

THEOREM 8.2. *If the surface S is referred to its asymptotic net as parametric,*

the normal to S at a non-parabolic point is the cusp axis of x with respect to the extremal curves of the integral invariant

$$\int (2K^{1/4}D'dudv)^{1/2},$$

where $2D'dudv$ is the second fundamental form of S and K is the total curvature of S at x.

References

V. G. Grove

1. *On canonical forms of differential equations*, Bulletin of the American Mathematical Society, vol. 36 (1930), pp. 582–586.

P. O. Bell

1. *Projective analogues of the congruence of normals*, American Journal of Mathematics, vol. 62 (1940), pp. 680–686.
2. *A study of curved surfaces by means of certain associated ruled surfaces*, these Transactions, vol. 46 (1939), pp. 389–409.

N. Abramescu

1. *Nouvelle méthode pour obtenir la cubique qui donne les tangentes de Darboux en un point d'une surface*, Comptes Rendus de l'Académie des Sciences, Paris, vol. 209 (1939), pp. 780–781.

E. J. Wilczynski

1. *Projective differential geometry of curved surfaces (first memoir)*, these Transactions, vol. 8 (1907), pp. 233–260.
2. *Projective differential geometry of curved surfaces (second memoir)*, these Transactions, vol. 9 (1908), pp. 79–120.

G. M. Green

1. *Memoir on the general theory of surfaces and rectilinear congruences*, these Transactions, vol. 20 (1919), pp. 79–153.
2. *Nets of space curves*, these Transactions, vol. 21 (1921), pp. 207–236.

Fubini and Čech

1. *Geometria Proiettiva Differenziale*, vol. 1, 1926; vol. 2, 1927; Bologna, Zanichelli.

E. P. Lane

1. *Projective Differential Geometry of Curves and Surfaces*, The University of Chicago Press, 1932.

University of Kansas,
Lawrence, Kans.

TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY

INDEX, VOLUMES 41–50

1937–1941

	Vol.	Pages
Adams, C. R., and Clarkson, J. A. The type of certain Borel sets in several Banach spaces	45_2	332–334
—— Correction to "Properties of functions $f(x, y)$ of bounded variation"	46_3	468
Adams, C. R., and Morse, A. P. On the space (BV)	42_2	194–205
—— Continuous additive functionals on the space (BV) and certain subspaces	48_1	82–100
Adkisson, V. W. Plane peanian continua with unique maps on the sphere and in the plane	44_1	58– 67
Agnew, R. P. Comparison of products of methods of summability	43_3	327–343
—— On kernels of faltung transformations	48_1	1– 20
Ahlfors, L. V. On Phragmén-Lindelöf's principle	41_1	1– 8
—— An extension of Schwarz's lemma	43_3	359–364
Albert, A. A. Symmetric and alternate matrices in an arbitrary field, I	43_3	386–436
Alexander, H. W. The role of the mean curvature in the immersion theory of surfaces	47_2	230–253
Alexandroff, P. General combinatorial topology	49_1	41–105
Ambrose, W. On measurable stochastic processes	47_1	66– 79
Baer, R. Groups with abelian central quotient group	44_3	357–386
—— Groups with preassigned central and central quotient group	44_3	387–412
—— Nets and groups	46_1	110–141
—— Nilpotent groups and their generalizations	47_3	393–434
—— Nets and groups. II	47_3	435–439
Beckenbach, E. F., and Reade, M. Generalizations to space of the Cauchy and Morera theorems	49_3	354–377
Bell, E. T. Polynomial approximations for elliptic functions.	44_1	47– 57
Bell, P. O. A study of curved surfaces by means of certain associated ruled surfaces	46_3	389–409
—— On differential geometry intrinsically connected with a surface element of projective arc length	50_3	529–547
Bernstein, B. A. Postulates for abelian groups and fields in terms of non-associative operations	43_1	1– 6
Berry, A. C. The accuracy of the Gaussian approximation to the sum of independent variates	49_1	122–136

	Vol.	Pages
Bird, M. T. On a classification of integral functions by means of certain invariant point properties. A supplement	42_2	191–193
Birkhoff, G. Analytical groups	43_1	61–101
Bliss, G. A. Normality and abnormality in the calculus of variations	43_3	365–376
—— Definitely self-adjoint boundary value problems	44_3	413–428
Boas, R. P. On a generalization of the Stieltjes moment problem	46_1	142–150
—— Expansions of analytic functions	48_3	467–487
Boas, R. P., and Widder, D. V. The iterated Stieltjes transform	45_1	1– 72
Bochner, S. Harmonic surfaces in Riemann metric	47_1	146–154
Brauer, R. On sets of matrices with coefficients in a division ring	49_3	502–548
Burington, R. S. On circavariant matrices and circa-equivalent networks	48_3	377–390
Busemann, H., and Mayer, W. On the foundations of calculus of variations	49_2	173–198
Calkin, J. W. Abstract symmetric boundary conditions	45_3	369–442
Cameron, R. H., and Martin, W. T. Analytic continuation of diagonals and Hadamard compositions of multiple power series	44_1	1– 7
Cameron, R. H., and Wiener, N. Convergence properties of analytic functions of Fourier-Stieltjes transforms	46_1	97–109
Camp, B. H. Correction to the paper "The multinomial solid and the chi test"	44_1	151
Carlitz, L. A class of polynomials	43_2	167–182
Chanler, J. H. The involution curve determined from a special pencil of n-ics	42_1	1– 15
Clarkson, J. A., and Adams, C. R. The type of certain Borel sets in several Banach spaces	45_2	322–334
—— Correction to "Properties of functions $f(x, y)$ of bounded variation"	46_3	468
Clifford, A. H., and MacLane, S. Factor-sets of a group in its abstract unit group	50_3	385–406
Coburn, N. Conformal unitary spaces	50_1	26– 39
Coolidge, J. L. Analytic systems of central conics in space	48_3	359–376
Copeland, A. H. Consistency of the conditions determining Kollektivs	42_3	333–357
Courant, R. On a generalized form of Plateau's problem	50_1	40– 47
Coxeter, H. S. M. The abstract groups $G^{m,n,p}$	45_1	73–150
Cramér, H. On the representation of a function by certain Fourier integrals	46_2	191–201

	Vol.	Pages
Curry, H. B. The paradox of Kleene and Rosser	50_3	454–516
Curtiss, J. H. On the Jacobi series	49_3	467–501
De Cicco, J. The geometry of whirl series	43_3	344–358
—— The differential geometry of series of lineal elements	46_3	348–361
—— The geometry of fields of lineal elements	47_2	207–229
De Cicco, J., and Kasner, E. Families of curves conformally equivalent to circles	49_3	378–391
Dilworth, R. P. Non-commutative residuated lattices	46_3	426–444
—— Ideals in Birkhoff lattices	49_3	325–353
Dilworth, R. P., and Ward, M. Residuated lattices	45_3	335–354
Doob, J. L. Stochastic processes depending on a continuous parameter	42_1	107–140
—— Stochastic processes with an integral-valued parameter	44_1	87–150
—— Regularity properties of certain families of chance variables	47_3	455–486
Douglas, J. A new special form of the linear element of a surface	48_1	101–116
—— Solution of the inverse problem of the calculus of variations	50_1	71–128
Duffin, R. J., and Schaeffer, A. C. A refinement of an inequality of the brothers Markoff	50_3	517–528
Dunford, N. Uniformity in linear spaces	44_2	305–356
Dunford, N., and Pettis, B. J. Linear operations on summable functions	47_3	323–392
Feller, W. On the integro-differential equations of purely discontinuous Markoff processes	48_3	488–515
Fialkow, A. Conformal geodesics	45_3	443–473
Fialkow, A., and Kasner, E. Geometry of dynamical trajectories at a point of equilibrium	41_2	314–320
Flood, M. M. Equivalence of pairs of matrices	44_1	8– 17
Franklin, P. A new class of transcendental numbers	42_2	155–182
Friedman, B. Analyticity of equilibrium figures of rotation	43_2	183–198
Friedrichs, K. On certain inequalities and characteristic value problems for analytic functions and for functions of two variables	41_3	321–364
Garabedian, H. L., and Wall, H. S. Hausdorff methods of summation and continued fractions	48_2	185–207
Gelbart, A. On the growth properties of a function of two complex variables given by its power series expansion	49_2	199–210
Gleyzal, A. Order types and structure of orders	48_3	451–466
Gore, G. D. Transformations of a surface bearing a family of asymptotic curves	43_2	303–320
Graustein, W. C. Extensions of the four-vertex theorem	41_1	9– 23

	Vol.	Pages
—— Harmonic minimal surfaces	47_2	173–206
Greville, T. N. E. Invariance of the admissibility of numbers under certain general types of transformations	46_3	410–425
Grove, V. G. Geometry of a surface in the neighborhood of a spine	45_2	256–264
Hall, D. W. On a decomposition of true cyclic elements	47_2	305–321
Hall, D. W., and Whyburn, G. T. Arc- and tree-preserving transformations	48_1	63– 71
Hall, M. An isomorphism between linear recurring sequences and algebraic rings	44_2	196–218
—— The position of the radical in an algebra	48_3	391–404
Halperin, I. On the transitivity of perspectivity in continuous geometries	44_3	537–562
Hamilton, O. H. Fixed points under transformations of continua which are not connected in Kleinen	44_1	18– 24
Hartman, P. Mean motions and almost periodic functions	46_1	66– 81
Hestenes, M. R. A direct sufficiency proof for the problem of Bolza in the calculus of variations	42_1	141–154
Hestenes, M. R., and McShane, E. J. A theorem on quadratic forms and its application in the calculus of variations	47_3	501–512
Hille, E. Contributions to the theory of Hermitian series. II. The representation problem	47_1	80– 94
Hollcroft, T. R. The binet of quadrics in S_3	42_1	32– 40
Hopkins, C. Concerning uniqueness-bases of finite groups with applications to p-groups of class 2	41_2	287–313
Ingraham, M. H., and Wolf, M. C. Relative linear sets and similarity of matrices whose elements belong to a division algebra	42_1	16– 31
Iyer, V. G. On effective sets of points in relation to integral functions	42_3	358–365
—— A correction to the paper "On effective sets of points in relation to integral functions"	43_3	494
Jackson, D. Orthogonal polynomials with auxiliary conditions	48_1	72– 81
Jacobson, N. Abstract derivation and Lie algebras	42_2	206–224
—— Restricted Lie algebras of characteristic p	50_1	15– 25
James, R. D. A problem in additive number theory	43_2	296–302
Jeffery, R. L. Functions defined by sequences of integrals and the inversion of approximate derived numbers	41_2	171–192
Jennings, S. A. The structure of the group ring of a p-group over a modular field	50_1	175–185

	Vol.	Pages
Jones, F. B. Concerning certain topologically flat spaces	42_1	53– 93
van Kampen, E. R., and Wintner, A. On the reduction of dynamical systems by means of parametrized invariant relations	44_2	168–195
Kasner, E. The geometry of isogonal and equi-tangential series	42_1	94–106
—— The two conformal invariants of fifth order	44_1	25– 31
—— Conformality in connection with functions of two complex variables	48_1	50– 62
Kasner, E., and De Cicco, J. Families of curves conformally equivalent to circles	49_3	378–391
Kasner, E., and Fialkow, A. Geometry of dynamical trajectories at a point of equilibrium	41_2	314–320
Kawaguchi, A. Geometry in an n-dimensional space with the arc length $s=\int\{A_i(x, x')x'^i+B(x, x')\}^{1/p}dt$	44_2	153–167
Ketchum, P. W. On the possible rate of growth of an analytic function	49_2	211–228
Kiokemeister, F. The parastrophic criterion for the factorization of primes	50_1	140–159
Klipple, E. G. Two-dimensional spaces in which there exist contiguous points	44_2	250–276
Kober, H. On a theorem of Schur and on fractional integrals of purely imaginary order	50_1	160–174
Langer, R. E. The boundary problem of an ordinary linear differential system in the complex domain	46_2	151–190
—— A correction to "The boundary problem of an ordinary linear differential system in the complex domain"	46_3	467
Laporte, O. Stereographic parameters and pseudo-minimal hypersurfaces, II	41_1	124–137
Lehmer, D. H. On the series for the partition function	43_2	271–295
—— On the remainders and convergence of the series for the partition function	46_3	362–373
Levinson, N. On the growth of analytic functions	43_2	240–257
Lewis, D. C. Contributions to the transformation theory of dynamics	46_3	374–388
Lewy, H. A priori limitations for solutions of Monge-Ampère equations. II	41_3	365–374
—— On differential geometry in the large, I. (Minkowski's problem)	43_2	258–270
—— Generalized integrals and differential equations	43_3	437–464
Loomis, L. H. The decomposition of meromorphic functions into rational functions of univalent functions	50_1	1– 14

	Vol.	Pages
Lorch, E. R. On a calculus of operators in reflexive vector spaces	45_2	217–234
—— The integral representation of weakly almost-periodic transformations in reflexive vector spaces	49_1	18– 40
Lubben, R. G. Concerning limiting sets in abstract spaces, II	43_3	482–493
—— Concerning the decomposition and amalgamation of points, upper semi-continuous collections, and topological extensions	49_3	410–466
McShane, E. J. Some existence theorems in the calculus of variations. I. The Dresden corner condition	44_3	429–438
—— Some existence theorems in the calculus of variations. II. Existence theorems for isoperimetric problems in the plane	44_3	439–453
—— Some existence theorems in the calculus of variations. III. Existence theorems for nonregular problems	45_1	151–171
—— Some existence theorems in the calculus of variations. IV. Isoperimetric problems in non-parametric form	45_2	173–196
—— Some existence theorems in the calculus of variations. V. The isoperimetric problem in parametric form	45_2	197–216
McShane, E. J., and Hestenes, M. R. A theorem on quadratic forms and its application in the calculus of variations	47_3	501–512
MacColl, L. A. Geometric aspects of relativistic dynamics	46_3	328–347
MacLane, S. The Schönemann-Eisenstein irreducibility criteria in terms of prime ideals	43_2	226–239
—— Steinitz field towers for modular fields	46_1	23– 45
MacLane, S., and Clifford, A. H. Factor-sets of a group in its abstract unit group	50_3	385–406
MacLane, S., and Schilling, O. F. G. Normal algebraic number fields	50_2	295–384
MacNeille, H. M. Partially ordered sets	42_3	416–460
Madow, W. G. Contributions to the theory of multivariate statistical analysis	44_3	454–495
Maker, P. T. Conditions on $u(x, y)$ and $v(x, y)$ necessary and sufficient for the regularity of $u+iv$	45_2	265–275
Marcinkiewicz, J., and Zygmund, A. On the behavior of trigonometric series and power series	50_3	407–453
Marden, M. Kakeya's problem on the zeros of the derivative of a polynomial	45_3	355–368
Martin, R. S. Minimal positive harmonic functions	49_1	137–172
Martin, V., and Roberts, J. H. Two-to-one transformations on 2-manifolds	49_1	1– 17
Martin, W. T. On a minimum problem in the theory of analytic functions of several variables	48_2	351–357

	Vol.	Pages
Martin, W. T., and Cameron, R. H. Analytic continuation of diagonals and Hadamard compositions of multiple power series	44_1	1– 7
Mayer, W., and Busemann, H. On the foundations of calculus of variations	49_2	173–198
Milgram, A. N. Decompositions and dimension of closed sets in R^n	43_3	465–481
Montgomery, D. Almost periodic transformation groups	42_2	322–332
Montgomery, D., and Zippin, L. Topological group foundations of rigid space geometry	48_1	21– 49
Moore, R. L. Concerning essential continua of condensation	42_1	41– 52
Morrey, C. B. On the solutions of quasi-linear elliptic partial differential equations	43_1	126–166
Morse, A. P. Convergence in variation and related topics	41_1	48– 83
—— A continuous function with no unilateral derivatives	44_3	496–507
Morse, A. P., and Adams, C. R. On the space (BV)	42_2	194–205
—— Continuous additive functionals on the space (BV) and certain subspaces	48_1	82–100
Murdoch, D. C. Structure of abelian quasi-groups	49_3	392–409
Murray, F. J. On complementary manifolds and projections in spaces L_p and l_p	41_1	138–152
—— Bilinear transformations in Hilbert space	45_3	474–507
Murray, F. J., and von Neumann, J. On rings of operators. II	41_2	208–248
Neuhaus, A. Products of normal semi-fields	49_1	106–121
von Neumann, J., and Murray, F. J. On rings of operators. II	41_2	208–248
von Neumann, J., and Schoenberg, I. J. Fourier integrals and metric geometry	50_2	226–251
Niven, I. Integers of quadratic fields as sums of squares	48_3	405–417
Oldenburger, R. Rational equivalence of a form to a sum of pth powers	44_2	219–249
—— Exponent trajectories in symbolic dynamics	46_3	453–466
Ore, O. On the theorem of Jordon-Hölder	41_2	266–275
—— On functions with bounded derivatives	43_2	321–326
Pall, G. On the arithmetic of quaternions	47_3	487–500
Perlis, S. Maximal orders in rational cyclic algebras of composite degree	46_1	82– 96
Pettis, B. J. On integration in vector spaces	44_2	277–304
Pettis, B. J., and Dunford, N. Linear operations on summable functions	47_3	323–392
Phillips, R. S. Integration in a convex linear topological space	47_1	114–145

	Vol.	Pages
—— On linear transformations	48_3	516–541
Pólya, G. Sur l'existence de fonctions entières satisfaisant à certaines conditions linéaires	50_1	129–139
Poritsky, H. Generalizations of the Gauss law of the spherical mean	43_2	199–225
Post, E. L. Polyadic groups	48_2	208–350
Price, G. B. The theory of integration	47_1	1– 50
Puckett, W. T. On 0-regular surface transformations	47_1	95–113
Radó, T., and Reichelderfer, P. A theory of absolutely continuous transformations in the plane	49_2	258–307
Randels, W. C. On the summability of Fourier series	41_1	24– 47
—— On an approximate functional equation of Paley	43_1	102–125
Raudenbush, H. W., and Ritt, J. F. Ideal theory and algebraic difference equations	46_3	445–452
Reade, M., and Beckenbach, E. F. Generalizations to space of the Cauchy and Morera theorems	49_3	354–377
Reichelderfer, P., and Radó, T. A theory of absolutely continuous transformations in the plane	49_2	258–307
Reid, W. T. A direct expansion proof of sufficient conditions for the non-parametric problem of Bolza	42_2	183–190
—— A system of ordinary linear differential equations with two-point boundary conditions	44_3	508–521
Rinehart, R. F. An interpretation of the index of inertia of the discriminant matrices of a linear associative algebra	46_3	307–327
Ritt, J. F. On a type of algebraic differential manifold	48_3	542–552
Ritt, J. F., and Raudenbush, H. W. Ideal theory and algebraic difference equations	46_3	445–452
Robbins, H. On the classification of the mappings of a 2-complex	49_2	308–324
Roberts, J. H., and Martin, V. Two-to-one transformations on 2-manifolds	49_1	1– 17
Saks, S. On some functionals. II	41_1	160–170
Schaeffer, A. C. Existence theorem for the flow of an ideal incompressible fluid in two dimensions	42_3	497–513
Schaeffer, A. C., and Duffin, R. J. A refinement of an inequality of the brothers Markoff	50_3	517–528
Schaeffer, A. C., and Szegö, G. Inequalities for harmonic polynomials in two and three dimensions	50_2	187–225
Schilling, O. F. G. Regular normal extensions over complete fields	47_3	440–454
Schilling, O. F. G., and MacLane, S. Normal algebraic number fields	50_2	295–384

	Vol.	Pages
Schoenberg, I. J. Metric spaces and positive definite functions	44_3	522–536
Schoenberg, I. J., and von Neumann, J. Fourier integrals and metric geometry	50_2	226–251
Schwid, N. A correction to "The asymptotic forms of the Hermite and Weber functions"	47_2	322
Scott, W. T., and Wall, H. S. A convergence theorem for continued fractions	47_1	155–172
Sewell, W. E. Generalized derivatives and approximation by polynomials	41_1	84–123
Sewell, W. E., and Walsh, J. L. On the degree of polynomial approximation to analytic functions: Problem β.	49_2	229–257
Sheffer, I. M. A correction	41_1	153–159
Shohat, J. On mechanical quadratures, in particular, with positive coefficients	42_3	461–496
Singer, J. A theorem in finite projective geometry and some applications to number theory	43_3	377–385
Sokolnikoff, E. S., and Sokolnikoff, I. S. Thermal stresses in elastic plates	45_2	235–255
Sokolnikoff, I. S., and Sokolnikoff, E. S. Thermal stresses in elastic plates	45_2	235–255
Spencer, D. C. On finitely mean valent functions. II	48_3	418–435
Stone, M. H. Applications of the theory of Boolean rings to general topology	41_3	375–481
Strodt, W. C. Irreducibile systems of algebraic differential equations	45_2	276–297
Synge, J. L. The absolute optical instrument	44_1	32– 46
Szász, O. Fourier series and mean moduli of continuity	42_3	366–395
—— On strong summability of Fourier series	48_1	117–125
Szegö, G. On the gradient of solid harmonic polynomials	47_1	51– 65
Szegö, G., and Schaeffer, A. C. Inequalities for harmonic polynomials in two and three dimensions	50_2	187–225
Tornheim, L. Integral sets of quaternion algebras over a function field	48_3	436–450
Torrance, C. C. Tangent lines and planes in topological spaces	41_2	193–207
Trjitzinsky, W. J. Theory of non-linear singular differential systems	42_2	225–321
—— General theory of singular integral equations with real kernels	46_2	202–279
—— Properties of growth for solutions of differential equations of dynamical type	50_2	252–294

	Vol.	Pages
Wald, A. Limits of a distribution function determined by absolute moments and inequalities satisfied by absolute moments	46_2	280–306
Wall, H. S. Continued fractions and totally monotone sequences	48_2	165–184
Wall, H. S., and Garabedian, H. L. Hausdorff methods of summation and continued fractions	48_2	185–207
Wall, H. S., and Scott, W. T. A convergence theorem for continued fractions	47_1	155–172
Walsh, J. L. On interpolation by functions analytic and bounded in a given region	46_1	46– 65
—— On the degree of convergence of sequences of rational functions	47_2	254–292
—— Note on the degree of convergence of sequences of analytic functions	47_2	293–304
Walsh, J. L., and Sewell, W. E. On the degree of polynomial approximation to analytic functions: Problem β	49_2	229–257
Ward, M. Linear divisibility sequences	41_2	276–286
—— The law of apparition of primes in a Lucasian sequence	44_1	68– 86
Ward, M., and Dilworth, R. P. Residuated lattices	45_3	335–354
Weyl, H. Theory of reduction for arithmetical equivalence	48_1	126–164
Whyburn, G. T., and Hall, D. W. Arc- and tree-preserving transformations	48_1	63– 71
Widder, D. V. The Stieltjes transform	43_1	7– 60
Widder, D. V., and Boas, R. P. The iterated Stieltjes transform	45_1	1– 72
Wiener, N., and Cameron, R. H. Convergence properties of analytic functions of Fourier-Stieltjes transforms	46_1	97–109
Williams, W. L. Permanent configurations in the problem of five bodies	44_3	563–579
Wintner, A., and van Kampen, E. R. On the reduction of dynamical systems by means of parametrized invariant relations	44_2	168–195
Wolf, M. C., and Ingraham, M. H. Relative linear sets and similarity of matrices whose elements belong to a division algebra	42_1	16– 31
Woodard, D. W. The characterization of the closed n-cell	42_3	396–415
Zariski, O. Generalized weight properties of the resultant of $n+1$ polynomials in n indeterminates	41_2	249–265
—— Pencils on an algebraic variety and a new proof of a theorem of Bertini	50_1	48– 70

	Vol.	Pages
Zippin, L., and Montgomery, D. Topological group foundations of rigid space geometry	48_1	21– 49
Zorn, M. Continuous groups and Schwarz' lemma	46_1	1– 22
Zuckerman, H. S. On the coefficients of certain modular forms belonging to subgroups of the modular group	45_2	298–321
Zygmund, A., and Marcinkiewicz, J. On the behavior of trigonometric series and power series	50_3	407–453
Errata, volume 41 (Morse, A. P. Convergence in variation and related topics, pp. 48–83; Ore, O. On the theorem of Jordon-Hölder, pp. 266–275)	41_3	482